Herding Behavior, Forecasts and Reactivity

Dissertation

submitted to the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to obtain the degree of

Doctor of Philosophy in Economics

presented by

Reto Cueni

from Kloten (ZH) and Dittingen (BL), Switzerland

approved in February 2014 at the request of

Prof. Dr. Dr. h.c. mult. Bruno S. Frey

Prof. Dr. Jean-Charles Rochet

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 12.02.2014

Chairman of the Doctoral Board: Prof. Dr. Josef Zweimüller

Contents

Figures

Tables

Preface

Had I not had the chance to attend Prof. Dr. Dr. h.c. mult. Bruno S. Frey's lectures as a master student, I would not have started the endeavor of writing a thesis in economics. I am most grateful that he inspired me to engage in writing this thesis and that he accepted me as one of his doctoral students. During the past five years at the University of Zurich's Department of Economics, he introduced me to a creative and unorthodox world of economic research, which always incorporated rigorous thinking and applied empirical work. From the very beginning, he took me and my fellow students along to research seminars and conferences and never hesitated to introduce us and our research to other scholars. He is not only a stimulating supervisor but also a supportive academic mentor. The research ideas for this thesis evolved from many discussions with him and with my colleagues Lasse Steiner, Jana Gallus, Vera Eichenauer, Stephan Kyburz, Prof. Dr. Susanne Neckermann, and Prof. Dr. Christine Benesch. I am particularly grateful to Lasse Steiner and Jana Gallus for their efforts in reading and commenting on earlier drafts of this thesis. Many thanks also go to Isabel Ellenberger, Lisa Elsasser, Martina Lauener, Rebecca Baumgartner and Jolanta Lindenmann for their secretarial support, which helped us in daily life to concentrate on our academic work. Prof. Dr. Jean-Charles Rochet kindly agreed to assess the thesis as a co-referee, for which I am very thankful.

The Chapters 2, 3 and 5 are based on papers jointly written with Bruno S. Frey; Chapter 4 is based on a paper co-authored with Prof. Dr. Peter Fiechter. To highlight their contribution, I use the first-person plural "we" throughout the whole thesis. All four chapters have been extended and comprise additional analyses and results. Of course, all remaining errors in this thesis remain my own.

I am grateful to Prof. Dr. Bruce Kogut and Prof. Dr. Stephan Meier for having invited me for a research semester at Columbia Business School and for their generous hospitality during this time. I also wish to thank Bruno S. Frey for providing me with this opportunity, as well as the Swiss National Science Foundation for awarding me a fellowship to afford my costs of living in New York City. In addition, I thank the Ecoscientia Foundation for their

research grant and Prof. Dr. Conrad Meyer, who acted as the contact person for the foundation at the University of Zurich.

For comments and suggestions on the papers and presentations on which this thesis is partly based, I am indebted to Prof. Dr. Christoph Engel, Prof. Dr. Kent Daniel, Prof. Dr. Alexander Wagner, Prof. Dr. Nick Netzer, Prof. Dr. Margit Osterloh, Prof. Dr. Hossam Zeitoun, Dr. Emil Innauen, Dr. Gregory Baetschman, Dr. Kevin Staub, Peter Rosenkranz, Sven Heinke, Patricia Feubli, Christian Ulbrich, Maximiliane Haut, everybody mentioned above and many other participants of research seminars at the University of Zurich, MPI Bonn, MPI Munich, Warwick Business School, IESE Business School, Humboldt University, ISNIE Florence and Columbia Business School. I also wish to thank my former colleagues Viktor Dammann, Patrick Rafaisz, and Remo Hartmann, who work or have worked as security analysts and who read Chapter 4 and commented on my work from a practitioner's perspective. Furthermore, I am very thankful to Jia Guo who shared her data on the Chinese art market with me, and to Ludwig Zgraggen, Christophe Voisard and Alexander Giordano for providing me with meteorological knowledge and data from MeteoSwiss. I also wish to thank the anonymous respondents who participated in the interview survey of Chapter 3, and Margit Osterloh, Bruno S. Frey, Sarah Camenisch and Remo Hartmann for helping me find the right interview partners. I thank Simon Milligan for the care with which he read the whole manuscript and improved the English.

Special thanks go to my friends and family who supported me during these years by listening to my stories of achievements and drawbacks. I also owe a debt of gratitude to Lasse Steiner, a steadfast fellow traveler through many long working days and also a friend for discussing the big questions in life for more than four years, and to Jana Gallus for academic and personal support during the last two years. I want to thank my parents who have always supported me and my endeavors and my brother, who taught me to bear the costs of non-herding. Finally, I am delighted that Simona Wanzenried has accompanied me on my way to accomplishing this thesis. She was and is wonderful in helping me keep the balance between all the many aspects of life and not lose my focus on the most important things.

Zurich, October 2013

1. Introduction

The question that led to this dissertation emerged in autumn 2009, a year after the collapse of Lehman Brothers and the outbreak of the latest financial crisis: Did nobody anticipate the crisis? And if there were people who did, why did they not advise caution based on their forecasts?

Indeed, several people issued public warnings about the imminence of a financial crisis. Economists such as Nouriel Roubini and Brad Setser (2004) or Paul Krugman (2007) stated before the emergence of the crisis that the amount of U.S. borrowing was dangerous, that the U.S. current account was unsustainable and that the dollar would collapse once this became evident to all.[1] Robert Shiller wrote in the preface of the second edition of his famous book *Irrational Exuberance* that he worried about the high discrepancy between the skyrocketing U.S. stock prices and the modestly moving earnings of the stocks, i.e. of the underlying companies (Shiller 2005, p. 4-14).

However, when analyzing the information flow before a crisis, it is important to identify how many of the academics' warnings reached the industry and whether there were also practitioners that drew attention to the critical situation. There are examples of such practitioners, most prominently Meredith Whitney, a formerly unknown analyst with Oppenheimer, a mid-sized U.S. brokerage house. Whitney produced a strongly negative report on Citibank that was issued on October 31, 2007, and led to an uproar in the market and a consecutive downgrade of Citibank by the majority of other analysts covering the stock. Whitney was able to raise her voice and to be noticed, which finally disciplined the herd of optimistic forecasters for Citigroup. After she guided the investors' attention to the toxic assets[2] on the bank's balance sheet, the credibility of analysts arguing the contrary shrank drastically (Hong and Kacperczyk 2010, p. 1685). So, why did more analysts not raise caution flags in the period before the crisis? Did only a few of them see the upcoming disaster?

The answer might lie in a less well-known but even more interesting case. In early 2007, Ivy Zelman, an analyst with the U.S. investment banking division of Credit Suisse,

[1] An extensive review of the literature is provided in Reinhart and Rogoff (2009, Part V.)

[2] The term *toxic assets* defines financial assets that have become illiquid, a condition which typically accompanies a strong decline in their value (see, e.g., Longstaff 2010).

started to downgrade the stocks of the construction and building supplier sectors she covered. In the following months, she was heavily criticized by both the managers of the downgraded companies and by sales persons from her own brokerage unit of the bank. Some sales representatives urged her to upgrade the stocks again and one even sent her an e-mail "warning that analysts who stayed bearish[3] too long often lost their job" (Robinson 2009, p. 33). At that moment, Zelman faced severe costs for not going with the herd. Had she followed the prevailing opinion of her colleagues and the CEOs of the companies she covered, it is likely that nobody would have asked her to justify her decision. "It was no fun being the bear", Ivy Zelman said two years later in an article in *Bloomberg Markets Magazine* (Robinson 2009, p. 34). In May 2007, she resigned from the bank to start her own independent brokerage house. When the "meltdown of the U.S. subprime market" began in summer 2007 (Reinhart and Rogoff 2009, p. 199), Zelman's call on the U.S. construction and building supplier sector turned out to be absolutely right. Economically speaking, Zelman faced major costs, from subtle psychological pressure to the threat of losing her job when she issued a forecast that was correct but went against the prevailing opinion in her organization and the market. These costs can be collectively termed the *costs of non-herding*.

What are the characteristics of such costs of non-herding and where do they come from? A first answer to these questions can be illustrated with an interview given in May 2011 by Brady Dougan, the CEO of Credit Suisse. Dougan stated in the interview that he had serious problems explaining to investors and even to some of his own colleagues why the bank's investment in the U.S. mortgage market had not been larger in spring 2007 (Binswanger and Gimes 2011, p. 18):

> "The crisis confirmed the, per se, simple insight that people tend to behave like a herd. That was before the outbreak of the financial crisis in 2008. Human nature is pro-cyclical and in particular in our business. Everybody wants to go there where everybody else sees success. In the year 2007, the head of our U.S. mortgage business was reproached from investors why he did not make as much money as Lehman Brothers[4]. That is the toughest task for a management: Not to lose one's

[3] Being *bearish* or *a bear* is financial markets' slang for expressing a pessimistic or cautionary view about the future of the markets, a stock, or a certain investment; it means the opposite of being *bullish* or *a bull*, i.e. expressing an optimistic view of the future.

[4] Before filing for bankruptcy on September 14, 2008, Lehman Brothers was the fifth largest investment bank in the world and one of the largest investors in the U.S. subprime mortgage sector, generating an extraordinary net revenue *increase* of 10% on the level of USD 19.3 billion, while Credit Suisse had to digest a *decline* of almost

> head in such a situation, keep on being critical and to convince the colleagues that it is the right way to hold off."

On the question if one has to go against the tide, Dougan responded:

> "First, the markets did not reward our change in the investment banking strategy. However, as in 2008 the crisis broke out, everybody asked to do business with us. […] Now it has been said that we are the best and the most secure bank, because we did not need any help from the government."[5]

Dougan's answer reveals several interesting features of the situation before the last financial crisis (and before a crisis in general) that can help to explain the absence of people raising their voices and advising caution. First, it is always more difficult not to go with the herd than to do what everyone else is doing. Second, the more successful a herd becomes during a certain period, the more will people who stay outside of the herd come under pressure to join the herd. Third, ex-ante, nobody knows if staying aside the herd will benefit the loners, but the latter have instant costs in justifying the action to others and to themselves. These features are central to the costs of non-herding and to the question of how these costs can induce the pressure on the individual to go with the herd. Surprisingly, as Dougan remarked, even powerful CEOs of large multinational companies face the costs of non-herding when deviating from the crowd.

It is important to make clear that deviating from the crowd is not always an ingenious action. In contrast, the "wisdom-of-crowds" can be an important tool for estimating a future outcome (Galton 1907; Lorge et al. 1958; Surowiecki 2005). Yet, to achieve a wisdom-of-crowds effect when aggregating the private information of individuals, the information has to be independent of the impact of others. In contrast, social influence distorts the efficiency of information aggregation. By providing private information to others, an individual exerts informational externalities and can bias their assessment of a future situation; but an individual might also exert a positive influence, as in the cases of the two analysts, Ivy Zelman and Meredith Whitney. On average, it has been shown by experiments that social influence tends to undermine the wisdom-of-crowds effect by reducing the diversity of

---

7% down to USD 18.9 billion in its investment banking business (see, e.g., Lehman Brothers Annual Report 2007; Credit Suisse Annual Report 2007; Fernando et al. 2012).

[5] The interview was released in German in a Swiss news magazine (*Das Magazin*). The interview excerpts have been translated by the author and cross-checked by an external copy-editor.

opinions in a group, resulting in biased information aggregation (Lorenz et al. 2011). Therefore, institutional factors are essential to keeping information providers in organizations and markets as independent as possible.

Finally, in reactive markets, in which opinions and forecasts can lead to self-fulfilling or self-defeating prophecies, distortions due to the costs of non-herding are exacerbated. In such markets, reactivity causes the problem that forecasts prompt people to change their behavior, and this change in behavior then influences the actual market outcome. As the forecasters try to predict the market outcome, they face the problem of forecasting a target that is moving endogenously due to their own forecasts. The self-fulfilling and self-defeating mechanism that a prophecy can trigger in any anthropogenic[6] system biases the incentives of the forecasters in two different ways (Morgenstern 1928; Merton 1936; 1948). Firstly, if it is assumed that a majority of congruent forecasts can cause the market participants to react in the predicted way, the herd of forecasters can induce a self-fulfilling prophecy. Under such circumstances, the forecasters that aim at staying outside the herd face even higher costs than without reactivity, as the herd’s forecast will become the actual outcome due to the market participants' reaction to it. Secondly, in the case of a self-defeating prophecy, for example the prediction of severe unemployment, the forecasters' incentives to issue a cautionary forecast are vitiated due to reactivity. The reactions of the market participants can ruin the forecast's ex post accuracy even though ex ante the forecast was correct. For instance, employees take measures to evade the predicted unemployment; they might accept wage cuts or similar actions. Thus, the self-defeating prophecy curbs the forecasters’ incentives to release such cautionary forecasts. The forecasters will anticipate low forecast accuracy when their predictions are assessed ex post by the market participants and, hence, will refrain from issuing cautionary forecasts even if it was ex ante correct.

The aims of this thesis are twofold. First, the thesis analyzes *institutional factors of organizations* that can help prevent herding behavior by reducing the costs of non-herding for individuals. The literature on herding behavior remains largely silent about institutional factors, such as an organization’s ownership, hierarchical structure, or wage policy, which might lower employees’ costs of non-herding. A comparative analysis is employed to identify useful institutional factors. Second, the problems for forecasters in *reactive* markets and the resulting distortions for the market participants are discussed from the perspective of various areas of the social sciences. A new institution for reactive markets is proposed that helps to

[6] Anthropogenic systems include all markets in which the actions of human beings influence the actual outcome.

reduce the pressure to herd and to provide a more balanced informational environment in such markets. Furthermore, a comparative analysis is conducted of markets with different levels of reactivity, and the influence of reactivity on forecast accuracy as well as on herding behavior is investigated empirically. The analysis exploits the difference in the level of reactivity in three markets: the market for weather forecasts, for financial forecasts, and for art price estimates. While in the case of the non-reactive weather the forecasts cannot influence the actual outcome, the forecasts in the cases of the reactive financial and art markets influence the final result.

The *effects* of herding behavior in organizations (e.g., Janis 1972; Strasser and Titus 1985; Scharfstein and Stein 1990; Garicano and Posner 2005) and on markets (Kindleberger 1978; Bikhchandani and Sharma 2001; Salganik et al. 2006; Akerlof and Shiller 2009) has been said to be detrimental to the efficiency of market outcomes and of decision processes in organizations. Therefore, insights into how to lower the pressure to herd by reducing the costs of non-herding for employees might interest managers that aim to improve the informational resources in their companies. In addition, regulators might be interested in reducing the costs of non-herding for employees of information providers in financial markets by insisting on the presence of important institutional factors of such organizations. Finally, investors and customers can reduce the problem of herding behavior in financial markets by carefully selecting banks or brokers with favorable institutional factors. In general, informing financial market participants about institutional determinants of banks and brokerage houses that lead to more unbiased and balanced information provision might help to prevent future crises (see, e.g., *The Economist* 2009). Furthermore, the study on reactivity is intended to draw attention to the exuberant assumption that more data and faster information processing will improve forecasts to full accuracy in anthropogenic markets. Additionally, the insights provided into the interplay between reactivity in markets and forecasters' reputation may raise the awareness of market participants about how reactivity can deteriorate the ex post assessment of forecasters’ accuracy.

The remainder of this thesis is organized as follows. Chapter 2 provides a literature review of the history of research on herding behavior across various fields from psychology to sociology to media research. The emphasis is placed on the social sciences and in particular on economics. Chapter 3 discusses how institutional and individual determinants of the costs of non-herding can help to reduce herding behavior in organizations. The theoretical propositions are reviewed with a qualitative analysis of twelve in-depth interviews with directors, highly ranked executives, and security analysts from various kinds of financial services companies. The interviewees make clear that an adjustment of certain institutional determinants should be able to reduce the costs of non-herding in organizations. They support the propositions of flatter hierarchies, less centralized structures, increased private ownership, a reduction of performance-related salary components and a more independent risk management. Chapter 4 quantitatively tests one of the aforementioned institutional factors, ownership structure, with a large data set of earnings estimates of security analysts. The data consists of more than one million observations between 1999 and 2008. The evidence indicates an economically and statistically significant effect of a brokerage's ownership structure on the propensity of its analysts to engage in herding behavior. Chapter 5 provides a theoretical introduction to reactivity and its effects on forecasting in a broad range of anthropogenic markets and non-human systems. In addition, the proposal of a new forecasting institution that acts as the devil's advocate in highly reactive markets is developed and its usefulness explained. Chapter 6 tests the propositions of Chapter 5 in a comparative institutional analysis comprising three forecast markets with different levels of reactivity. The analysis employs forecast data from the non-reactive weather system, in which the forecasts do not influence the actual outcome, and data from the reactive art and financial markets, in which the forecasts influence the outcome they try to predict. The study delivers empirical evidence that the vast increase in data availability and information-processing technology does not improve the forecast accuracy for anthropogenic market outcomes. In contrast, in non-human systems such as weather, enhanced data availability and faster processing techniques have increased the forecast accuracy significantly. In addition, the evidence presented in Chapter 6 shows that the level of reactivity influences the herding behavior of forecasters in anthropogenic markets and that forecasters with a high reputation can use their reputation to induce biased herding forecasts. Finally, Chapter 7 concludes and critically examines the results and the limitations of the studies provided, reconsidering the implications for economics, market participants and regulators and suggesting directions for future research.

2. Literature on Herding Behavior and Related Concepts[7]

The scientific literature on herding behavior is immense. Every area of the social sciences bears treatises on the topic, sometimes using closely related terms such as conformity, mass psychology, or group or collective behavior. Even in sociobiology or in fields such as computer science, the topic is intensively scrutinized. Yet, each of the individual areas are mostly silent about the research done in other fields. This literature review aims at combining and integrating the various areas that treat the topic of herding behavior in the field of the social sciences, focusing on the area of economics. To reasonably perform this task, only the most important milestones can be considered. However, the literature review provides an extensive interdisciplinary overview on the topic of herding behavior, organized in chronological order for a better understanding of the gradual development in the various areas of the social sciences.

Two of the earliest contributors in the research on herding behavior and the psychology of the crowd were the French sociologist Le Bon (1895) and the political economist Veblen (1899).[8] In his book *La Psychologie des Foules*, Le Bon describes the sociopolitical dangers and risks evoked by human herding behavior. He defines the "crowd" from a psychological point of view, stating that "the crowd is always dominated by considerations of which it is unconscious" (Le Bon 1895, p. 6). In contrast, the economist Veblen, in his book *The Theory of the Leisure Class,* analyzes the herding behavior of different social classes with regard to their consumption. The various "social herds" try to distinguish themselves by consuming similar products which they know that a lower social class cannot afford to buy or to use. Veblen uses the example of a woman's dress that is so conspicuously voluminous and impractical that it should be avoided and seen as pure waste. Yet, a social class which has

[7] This chapter draws on two papers written jointly with Bruno S. Frey (Cueni and Frey 2012; forthcoming).

[8] Of course, there exist earlier *descriptions* of herding behavior in markets, one of the most prominent one being Mackay's "Extraordinary Popular Delusions and the Madness of Crowds" (1841). Mackay describes the frenzy in the market for tulips in the Netherlands between 1634 and 1636, the peak of the tulip bubble. However, these treatises do not analyze but rather merely describe the incidence of herding behavior. Garber (2000) provides a critical analysis of Mackay's work and an insightful essay of early price bubbles that misleadingly have been attributed to frenzies due to wrong or wrongly interpreted records of older works and chronicles.

enough economic resources to invest in such dresses and, hence, to render their wearers unable to perform any manual labor is using the dresses to distinguish itself from another class that cannot afford to do so. From this economic perspective, individuals are conscious of their actions in forming a herd, as they derive utility from doing so, and are not unconscious as in Le Bon's treatise.

After the Second World War, scholars turned again to research on human herding behavior. In the fifties, Leibenstein (1950) takes up Veblen's (1899) idea on mass consumption and develops the theory of the *bandwagon effect,* due to which people tend to go along with what others do or think although they as individuals would act or think differently. In other words, they jump on the "bandwagon"[9] (Leibenstein 1950, p. 184).

At around the same time, Asch (1951) provided the first experimental evidence on the herding behavior of individuals with his seminal group experiments and thus invented a new chapter in sociopsychological research in this field. In Asch's experiments, there was usually a group of eight male students consisting of an "instructed majority" of seven individuals and a single "critical subject" (Asch 1956, p. 3-4). The instructed majority sometimes provided obviously wrong answers so as to create a situation in which the critical subject, who was not informed about the acting of the instructed majority, was forced to disagree when he intended to give an objectively correct answer. The critical subject was mostly placed as second last respondent of the whole group. The experiments showed that individuals (critical subjects) conform even to the obviously incorrect opinions of a majority. Asch inferred that social pressure due to the face-to-face communication in the experimental sessions triggered the participants' herding behavior. However, Deutsch and Gerard (1955) re-conducted Asch's experiments but eliminated the face-to-face contact between participants. They demonstrated that it was the informational effect of prior responses rather than Asch's social-pressure effect which was the dominating force behind the herding of the participants.

Later research integrated this idea and showed that the critical subjects' conformity with the majority could be rationalized by the subjects' assumption that all the other participants had no incentive to lie. Thus, following this assumption, the probability of seven wrong answers in a row – although the correct answer seems so obvious – was perceived to be very low. Paradoxically, this calculation led to a different explanation of Asch's findings. The

[9] A bandwagon was at that time a marketing instrument. It was a wagon on which a band was placed to play songs and which was pulled through the streets to attract peoples' attention to the banners on and around the wagon. In particular, this form of advertisement was used during political campaigns; "to jump on the bandwagon" means to join a certain political party or to support a candidate.

participants did not herd due to the social pressure to conform but due to their rational inference that they must have missed an important fact in the experimental situation or that they have been subject to an optical illusion, since the experiments were visual tasks. Hence, one can argue that the participants exhibited the high level of herding behavior in Asch's experiment *because* the answer of the instructed majority was so obviously wrong and not *although* it was so wrong. The subjects' conformity is mostly not due to pressure from the instructed majority but due to the rational analysis by the critical subject that the answer of the other seven must be true, as any other explanation seems to have a very low probability (Shiller 1995, p. 182-183). Indeed, if Asch had not told the seven members of the instructed majority to give the wrong answer, it is credible to assume that they would never have provided such wrong answers, or at least not so many of them. Indeed, in a control treatment without an instructed majority, more than 99% of all answers were correct. From this point of view, Asch's experiment could be seen as an example of the credulousness of individuals in not anticipating the biased information provision of the other participants due to distorted incentives rather than a case of conformity induced by social pressure.

Both Festinger (1954) in his social comparison theory and later Bandura (1965; 1976) with the social learning theory developed positive views of herding behavior, since imitating can also lead to learning; today, economists use social learning theory in their work on herd behavior (see, e.g., Hirshleifer and Teoh 2003). The difficulty is that social learning can lead to positive outcomes when the correct actions of others are imitated, or to negative outcomes when the wrong actions are reproduced. This difference is important for research on herding behavior, as the observation that people follow others in their actions can imply an inefficient imitation or an efficient adaptation. While the term herd behavior usually bears on the first, negative case, social learning usually refers to the latter, positive one.

In sociology, Granovetter (1973) developed a threshold model explaining the non-linear characteristics of herd behavior such that there exists a threshold of net costs to the individual above which the individual is no longer willing to bear the costs of not going with the herd. Hence, costs exceeding such a threshold trigger the adoption of new trends. Below the threshold, no such adoption takes place, not even gradually. Burt (1982) advanced research on herding behavior by studying peer groups. According to his theory of social networks (based on Festinger's theory of cognitive dissonance 1957), individuals tend to compare themselves with others, who are similar to them but differ in the factor in which they are interested. If such similar others (i.e. peers) engage in identical actions, similar individuals tend to follow a herd of their peers. According to Burt's theory, individuals do not seek to follow a random person; instead, they follow a definable "group of similar others" (Marsden and Friedkin 1993, p. 129). Based on this literature, three social-psychological concepts emerge that are important for the study of the herding behavior of employees in organizations.

Janis (1972) formulated his groupthink theory of conformist behavior during group deliberation processes, and it is a concept that is still in use (Baron 2005; Bénabou 2013). It suggests that participants in group decision processes miss important alternatives because they do not deliberate or discuss the "usual" solutions critically enough because they are pressured to conform so as to avoid conflict among the group members. Janis (1972, p. 209-216) describes some institutional factors which can be used to lower the pressure to conform on individuals in decision groups, for example establishing several independent groups working on the same problem, or assigning the role of critical evaluator to every team member.

Based on the problems of groupthink, the field of diversity research started to evolve during the eighties and provided a different perspective on herding behavior: more diverse teams, with respect to gender, education and social background, were assumed to make better decisions in crisis situations than more homogeneous teams (Bantel and Jackson 1989). Early empirical surveys supported this view, suggesting that members of more diverse teams process similar information in more diverging ways and, hence, were less prone to herd behavior. Subsequent studies, however, challenged this view and took a more critical approach to diversity and its impact on conformity and herd behavior. In particular, they provide evidence that mere diversity is not key to better group outcomes and that the context and conditions under which diversity is favorable remain unclear (Ely 2004).

As early as the seventies, Noelle-Neumann's (1974; 1984) theory of the spiral of silence described a further phenomenon in opinion research: In a dynamic process, various players start to adopt the opinion they assume to be the future majority opinion, even though personally they may have contrary views. According to Noelle-Neumann's explanation, the players do so because they try to avoid conflict and do not want to feel isolated and perceived as mavericks. This theory has been applied to financial markets (Aspara et al. 2008), as well as to organizational theory (Bowen and Blackmon 2003).

From an economic perspective, herding behavior can be studied by analyzing payoffs and network externalities. Interestingly, a first example of an analysis of a payoff interaction between rational and truly selfish individuals stems from sociobiology. Hamilton's (1971) study of selfish herds reveals that the clumping of animals is an indirect outcome of the selfish behavior of each animal in trying to put another member of the herd between itself and a predator. As in the case of decision making in teams, individuals avoid being outside the herd. As long as an individual goes along with the opinion of the majority, he or she can hide in the herd. In particular, if the decision turns out to be wrong ex post, the individual shares the blame with all others. This herd behavior serves as an individual risk-reducing strategy that involves negative externalities to the other individuals in the organization.

In the late eighties, the topics of herding behavior and resulting negative externalities were revisited. An analysis of lock-in effects showed that a whole industry could adopt certain standards although they were not the best ones in either technical or economic terms (Katz and Shapiro 1985; David 1985; Frank and Cook 1988; Arthur 1989). Due to the negative externalities exerted by the firms that follow the (suboptimal) standard, the other firms are also pushed to adopt the same standard; if they refuse to, they face various problems in the sales markets. For example, their sound storage medium will not fit a widely used music player due to different standards.

In the nineties, Scharfstein and Stein (1990) as well as Banerjee (1992) and Bikhchandani et al. (1992) advanced the topic of herding behavior in economics.[10] Banerjee and Bikhchandani et al. explain the herding behavior by information cascades, in which rational agents follow the strategy of the first mover in a sequential game and ignore their own private information. As a result, *information-based herding* can occur although the players have contradictory private information. In contrast, Scharfstein and Stein's model follows the logic of principal-agent problems and, hence, describes a *reputation-based herding* of employees.

Firstly, models based on *information cascades* nicely show the negative externalities of herding behavior. After randomly occurring, congruent actions by two or more successive players, the cascade begins to evolve and the players' private information is no longer taken into account in their decision process (Banerjee 1992; Bikhchandani et al. 1992). Such information-based herding-models have been amended by introducing information costs, imperfect information about the decision process of the other players, heterogeneous possibilities in decision making, and adaptive markets (see, e.g., Bikhchandani et al. 1998; Avery and Zemsky 1998). However, this strand of the economic literature is important as

[10] See also Chamley's (2004) extensive overview on herding behavior and social learning that focuses on theoretical models of the phenomenon.

informational cascades demonstrate how information can be lost. This happens because the first mover in a cascade makes a decision randomly, whereas the following movers ignore their private signals and imitate the decision of the first mover, assuming that he or she had better information than they did. Kuran and Sunstein (1999) show how *availability cascades* can have a similar effect to that of information cascades. For example, availability cascades related to the publicity of an investment opportunity reveal that the larger the amount of favorable publicity (without any new information), the more funds can be raised for an investment project, although there is no objective reason to favor this specific project over another.[11] Kuran and Sunstein also show how availability cascades occur in the political process of opinion formation and similar herding processes.

Secondly, *reputation-based* models connect the mechanism of hiding in the herd and the agency problems between supervisors and subordinates or clients and managers. The strategic actions of employees, who want to push their careers or gain a higher reputation, can evoke harmful herding behavior in organizations (Scharfstein and Stein 1990; Prendergast 1993; Zwiebel 1995).[12] Scharfstein and Stein provide a seminal model that considers two managers whose abilities are not known by any observers (e.g., clients or supervisors). Instead, observers infer the managers' abilities by comparing the investment decisions of the managers. After each period, the observers update their information according to the investment payoffs from the managers' strategies. The managers are paid based on the observers' assessments of their abilities. The outcome of the model is always that one manager follows the other's lead. The strategy thus results in a loss of information, because the manager that follows ignores his or her private information. As the observers do not know, ex ante and ex post, what they should expect as the best investment strategy, they compare the two managers. As long as the managers take the same action, neither of them is blamed for a bad payoff. The managers act rationally and use the herd as a risk-sharing mechanism.

[11] As attention and not only availability is included in the concept, it is also called *attention cascades* in the literature (see, e.g., Hirshleifer and Teoh 2010, p. 159; Kim and Meschke 2013, p. 1)

[12] Here we do not think of a harmful action inside the organization, which might be subject to legal issues and would make the case for whistleblowing and not only a case of non-herding or speaking up. Although whistleblowing is a closely related topic, we do not consider it here. We follow the argument of Premeaux and Bedeian (2003, p. 1538) and define speaking up and non-herding as evolving "from a desire to improve an organization by suggesting different approaches" rather than evolving from "perceived violations of personal principles," as in the case of whistleblowing (Miceli et al. 2008).

Zwiebel's (1995) extension makes the reputational herding model more realistic because in his version it is not always better to fail conventionally than to succeed unconventionally.[13] This means that the benchmarking of the managers' performances allows the managers of mediocre ability to hide in the herd, but the most able managers still have an incentive to deviate and build up an outstanding reputation. Similarly, Prendergast (1993) and his theory about "yes men" directly explore how conformity can emerge when firms use subjective performance evaluations. Instead of collecting and communicating important information about a project, workers try to elicit and represent their supervisors' opinion in order to please them and to further their careers.

Research by Graham (1999) and Hong et al. (2000) empirically supports the existence of reputation-based herding of financial analysts and reveals the different incentives at work depending on the stage of the analyst's career and other individual determinants. For example, they provide evidence that younger analysts herd more because they lack reputation. Similarly, Chevalier and Ellison (1999) show that the same holds for portfolio managers.

The literature includes many other empirical studies focusing on herding behavior in financial markets (see, e.g., Lakonishok et al. 1992; Christie and Huang 1995; Grinblatt et al. 1995; Nofsinger and Sias 1999; Wermers 1995). However, devising empirical tests to reveal the existence of truly inefficient and harmful herding might be "easier said than done and may even be impossible" (Bikhchandani and Sharma 2001, p. 281). It is very difficult to distinguish harmful, i.e. *intentional* herding, from *spurious* herding, which evolves due to new information in the market and which shows a congruent rational adaptation of market participants to the new information, leading to a "wise" herd (see Bikhchandani and Sharma 2001; Hirshleifer and Hong Teoh 2003; 2009).[14]

Experimental studies were partially able to show the herding behavior of market participants. However, the experimental testing of the different models showed that herding behavior can be explained with both information-based (Drehmann et al. 2005) and reputation-based models (Hey and Morone 2004), whereas the more realistic experiments favored the latter models (Cipriani and Guarino 2005). Alevy et al. (2007) provide evidence from a field experiment that under specific circumstances professionals are less prone to information cascades than inexperienced persons (e.g., students).

[13] This dictum was coined by John Maynard Keynes (1936, p. 158).

[14] *Spurious* indicates the misleading attribution of an observed conformity of individual actions to a herding behavior. The individuals did not depend their actions on the actions of others but inferred independently how to react upon certain new information in the market and come (non-intentionally) to the same conclusion.

Closely related to the topic of herding behavior are several sociopsychological surveys approaching the subject via *hidden profiles*. Starting with Strasser and Titus (1985), this strand of the literature tries to unveil the decision processes in groups: how information is shared and aggregated. At the beginning of the experiment, each member of the group receives only a piece of the whole information that the group needs to solve a problem. During the experiment, the researchers analyze how the various members of the group contribute their private information and how the group finally finds the proper solution or why it does not. As every group as a whole has all information necessary to find the proper solution, it is interesting to see that many of them fail to aggregate the information properly. Some members do not want to come into conflict with others and therefore do not contribute their contradictory but important information to the group decision process. The results of these experiments support the notion that non-herding is accompanied by costs and illustrate the resulting problems for decision processes in organizations (see, e.g., Schulz-Hardt et al. 2006).

The newer literature on management studies and organizational science focusing on *voice* and *silence* in organizations takes up many of these sociopsychological issues. This literature uses voice to define a situation in which the employee speaks up against a prevailing opinion and does not follow the herd, while silence is used as a situation in which the employee goes with the herd. In their seminal paper on silence in organizations, Morrison and Milliken (2000) touch on Janis's (1972) idea of groupthink and the research on diversity in management teams (Bantel and Jackson 1989). In a special edition of the *Journal of Management Studies* on the topic of the dynamics of voice and silence in organizations, some authors refer to Noelle-Neumann's (1974) spiral of silence, too (see, e.g., Bowen and Blackmon 2003). They use Noelle-Neumann's theory to describe the dynamics of silence in organizations or, in other words, the evolution of herding behavior within an organization.

Throughout the literature on voice and silence in organizations, Hirschman's (1970) concept of *exit*, *voice*, and *loyalty* serves as a general framework. The terms are used to describe the various alternatives of action of an organizational member during a decision process. The member can exit the organization, raise his or her voice against a certain decision, or remain loyal and follow the decision taken. What is particularly interesting for our purposes is that the empirical literature reveals valuable insights on the individual determinants of employees' probability of using their voices in an organization (see, e.g., Withey and Cooper 1989;

Tangirala and Ramanujam 2008). Hirschman's concept has been adjusted and modified to today's view of voice and silence in organizations.[15]

In 2001, Banerjee and Somanathan wrote their paper "A Simple Model of Voice" and completed the circle that started with Banerjee's (1992) "Simple Model of Herd Behavior". Paraphrasing Hirschman's concept of voice in organizations, Banerjee and Somanathan provide a model that relies on reputation-based herding without stating it explicitly.[16] Members of a group can contribute more or less private information to the decision-making process and thereby exert more or less influence on the decisions taken by the leaders. The decision about the amount of private information to contribute is based on the costs of their communication to the members of the group and the costs of reaching the leader's ear. In the model, these costs are defined by the level of homogeneity of opinions among the group members and by the level of difference in opinion between a member and a leader. In the most simple case of the model, moderation drives out extremism, as group members with extreme opinions always communicate their private information and, thus, neutralize each other. However, most importantly for our study, the authors are able to model a situation in which a leader with extreme opinions only receives extreme information which is close to his or her own extreme point of view. This is due to the fact that the costs of communication for group members with an opinion highly contrary to the extreme opinion of a leader exceed the benefits from it. These group members, who in the most simple case of the model would communicate their private information and move an extreme leader towards a more moderate position, remain silent in the most realistic version of the model. The mechanism that leads to this result is triggered by the realistic assumption that all members of the group can lie and that therefore a leader discounts the group members' information, and discounts them the more, the further away they are from his or her own opinion. Such a mechanism provokes an increase in extreme opinions and a reduction of information flow in the organization, even though it would be most valuable, in precisely those situations involving leaders with extreme opinions, to balance the different views.

The outcome of Banerjee and Somanathan's most realistic model is in line with the qualitative studies by Ryan and Oestreich (1991) or Milliken et al. (2003) on employee silence. They show that employees often fear to speak up and tend to herd, in particular upwards (to their

[15] See, for example, how Premeaux and Bedeian (2003) differentiate between Hirschman's (1970) notion of voice and their own expression of *speaking up*, which takes into account only that an employee uses his or her voice to make an improvement and not to express dissatisfaction with an organizational issue.
[16] In contrast, Banerjee's (1992) model focused on information-based herding.

supervisors), because they figure that they face negative consequences which do not depend on the true content of the information but on the content as perceived by the supervisor. Paradoxically, leaders with extreme opinions drive out moderate or controversial information that would be most valuable to them, increasing herding behavior in their organizations and losing information for the decision-making process.

In each of the theories reviewed, the costs of the individuals that are not going with the herd form a crucial factor which can be used when building a framework to further analyze the determinants of herding behavior. In addition, the theories support the notion of asymmetrical characteristics of the costs of non-herding, such as that every individual going with the herd can share the blame if the herd was proven wrong, but an individual who opposes bears the costs individually. The framework of the costs of non-herding and their determinants will be extensively discussed and qualitatively analyzed in the next chapter.

3. The Costs of Non-Herding in Organizations – An Interview Study[17]

3.1. Introduction

Aggregating the knowledge of many people is a powerful mechanism. When the employees of a firm do not provide their private information to the decision-making process in the firm, the decisions taken will be inefficient, as information will be lost. Yet, every employee that contributes information to an organization's information pool faces the risk of getting into trouble because some other employee might be offended and could react harshly. Economically speaking, employees face various kinds of potential costs if they choose to reveal and communicate their private information in an organization. These costs can effectively silence the employee. For example:

> "I raised a concern about some policies and I was told to shut up and that I was becoming a trouble maker. I would have pursued [the issue] further but presently I can't afford to risk my job. This has made me go into a detached mode, making of me a 'yes man' (male respondent, Information Systems)."[18]

This mechanism leads to two problems in organizations. First, the more employees remain silent, the weaker the information content of the decision-making process becomes. Second, the more employees remain silent and fail to criticize the prevailing opinion in an organization, the more the dominant opinion is perceived to be the correct one. In a dynamic process, this silence in organizations leads to herding behavior because the prevailing opinion becomes increasingly entrenched, and employees acquiesce so as not to be perceived as mavericks or troublemakers (Morrison and Milliken 2000; 2003). This dynamic can be seen as a risk-sharing mechanism (Hamilton 1971; Scharfstein and Stein 1990); expressing one's own beliefs and opinions always bears the risk of being personally offended, while sharing

[17] This chapter draws on two papers written jointly with Bruno S. Frey (Cueni and Frey 2012; forthcoming).
[18] Quoted from Milliken et al. (2003, p. 1453). See also the literature review in the preceding chapter for models about "yes men" behavior due to reputation-based herd behavior in organizations (Prendergast 1993).

the prevailing opinion in an organization splits the risk among all the followers. Furthermore, if the individual statement leads to a change in the decision-making process, the employee has to bear the responsibility for the decision. Economically speaking, if an employee deviates from the prevailing opinion, although the deviating input might have value for the whole group or organization, the employee has to bear the cost of not following the herd alone. Therefore, the costs of non-herding (CONH) are the employee's costs associated with expressing an opinion or having a voice in an organization. By lowering the CONH and therefore fostering the employee's participation with his or her private information, the organization can aggregate more information and improve its decision-making processes.

This chapter analyzes the determinants and problems of herding behavior in organizations and uses the framework of the CONH to understand the underlying mechanisms. The next section discusses these costs of non-herding extensively and formulates a first set of general propositions. Section 3.3 considers options for diminishing the CONH. Therefore, it focuses on the institutional and individual determinants of the CONH and formulates empirically testable propositions. Section 3.4 provides an explanation of the empirical method of the qualitative interview study with chairpersons and high ranked executives working in the financial industry and the resulting data evaluation. Section 3.5 challenges the propositions by comparing them to the interview results, and the last section concludes.

3.2. The Costs of Non-Herding

Whenever several people make a joint decision, costs of non-herding (CONH) arise. In a typical situation, one could imagine a bank's investment committee consisting of eight members. Assume that these members are equally ranked (later, we also discuss the case when the members have different hierarchy levels). These eight members of the committee have to decide about a specific investment, with one of them being in charge of the investment project. This manager makes a presentation about the anticipated opportunities and threats of the investment. Subsequently, the usual question-and-answer round starts, but nobody asks any crucial questions; the questions only scratch the surface of the possible problems. After hearing the answers to the harmless questions, the committee votes on the investment and unanimously adopts it. All of this happens although the investment project is not at all above suspicion.

Why do the committee members not ask more critical questions about the project? The individual members of the committee have several reasons not to raise any fundamental issue: Nobody wants to prolong the meeting; nobody wants to raise his or her voice against the project because an explanation would be required; nobody criticizes the project because they hope that his or her own project will also be approved smoothly; nobody wants to be the loner, and as long as no other member of the committee speaks up all the members keep quiet. In addition, because this silence has occurred several times before, nobody feels like defying the investment project this time, and thus everybody continues to herd.

These are only a few examples of the possible costs members of committees could face during a decision-making process (see, for more examples in various real-life contexts, e.g., Janis 1972; Garicano and Posner 2005). We want to show, however, that such situations reveal typical characteristics and mechanisms that can be identified and positively influenced. First, if committee members use voice, they face the CONH immediately. In contrast, the possible benefits lie in the future and can only be calculated after the project is realized. The member does not know in advance whether he or she will ever gain any benefit from it, while the costs are imminent. This shows four types of asymmetries. The first of these is in the temporal dimension between short-term costs and long-term benefits, and the second is in the differing probabilities of the costs and benefits occurring; the costs are rather clear, whereas benefits are very uncertain and difficult to calculate. Third, the costs have to be borne alone by the individual member of the committee, at least at an early stage, until other members join in the critique. The critic cannot benefit from a cost- and risk-sharing mechanism and hide in the herd. Fourth, there exists an asymmetric dynamic, as the costs are self-reinforcing. If the critics do not use voice in the first meeting, they face higher costs of doing so in the second meeting and so forth; in addition, the lower the number of the critics in a group, the higher are the critics' costs.

If we add the element of different ranks among committee members, principal-agent problems have to be considered. Assume that the supervisor of a committee member presents an investment project. To argue against this project becomes even more costly for the subordinate member. The supervisor may have influence or control over the future promotions of the inferior and over his or her future pay. These examples illustrate that any dependence on a supervising person can influence the costs of not herding dramatically for a specific committee member (e.g., Prendergast 1993).

At the organizational level, the problem of the CONH is that they induce other members to imitate behavior. By doing so, the individual members do not contribute their private information during the decision-making process, and so they vote for decisions which are not in accordance with their personal evaluations. This imitative behavior results in a loss of information, which is the major cost of herding behavior for the organization. In addition, the CONH trigger a self-reinforcing process. The longer the committee members do not raise any issues, the higher are the CONH and the lower are the incentives to speak up. The CONH create a vicious circle which can harm the organization in several ways. Chosen paths cannot be changed, and new and better investments or projects are missed or ignored. As decisions often involve many steps and are designed as a bottom-up process, there are CONH at each level that reinforce the herding behavior in the organization. Therefore, organizations should try to reduce the CONH.

3.2.1. Different Sources of Costs of Non-Herding

The CONH can arise from various sources, and can affect all employees. We pool these different sources into three categories: costs due to personal factors, costs due to conflicts between equally ranked employees, and costs between employees and their superiors.

Depending on self-perception, role behavior, and identity, the CONH can affect the individual employee in different ways (see, e.g., Withey and Cooper 1989; Morrison and Milliken 2000; Premeaux and Bedeian 2003): People who like to expose themselves or see their role as the devil's advocate perceive the CONH to be lower than people without such a character trait, for example. The employee's perceived costs of speaking up might also be influenced by individual determinants such as age, tenure with the company, experience in the job, or level of education.

Another source of the CONH lies in the relationship between the employee and the employee's coworkers or peers. Members of teams or committees usually experience a pressure to conform, and the members create their own social identity (Tajfel and Turner 1986). The more homogeneous a team is with respect to important individual aspects (professional background, education, gender, etc.), the more likely it is that groupthink occurs (Janis 1972; Bantel and Jackson 1989; Withey and Cooper 1989). These diverse phenomena of social interactions in teams influence the individual perception of the CONH. The more conformity is needed to be accepted as a member of the team, the higher are the costs if the individual employee deviates from the prevailing opinion. Employees are usually in competition with peers, which can reduce the CONH as it provides an incentive for the employees to speak up and distinguish themselves. In contrast, competition can also raise the CONH if the employee and his or her peers are rated by a superior manager using a benchmark (see Scharfstein and Stein 1990).

The last source of the CONH stems from the character of principal-agent relationships. In addition to the sources mentioned above, the relationship between the employee and superior, their personalities, experiences, and views on leadership have a massive influence on the CONH (see, e.g., Ryan and Oestreich 1991; Milliken et al. 2003). If the superior has to evaluate the employee's performance, and if this evaluation determines bonus payments, future project assignments, or promotions, the employee's CONH increase strongly. Due to career and reputational concerns, the employee's incentive to contradict the supervisor on a project is likely to vanish. If the superior is accorded much respect in the committee due to an impressive track record and extensive experience, the employee's CONH increase even further. These sources spawn a wide range of CONH for the employee, from small hostilities to bullying and sidelining, to a career-ending transfer of the employee, or even to dismissal. In summary, the costs of non-herding are characterized by

Characteristic C1: costs according to conflicts that occur before a decision is taken;

Characteristic C2: costs of justification after a decision is made;

Characteristic C3: temporal asymmetry, as large parts of the costs accrue in the short term and before a decision is made, while the benefits of non-herding only materialize in the long run, after a decision is made;

Characteristic C4: asymmetry in probabilities, as the costs are usually highly probable while the benefits are more uncertain;

Characteristic C5: social asymmetry, as costs have to be borne individually while benefits are mostly shared with other members of the organization;

Characteristic C6: dynamic asymmetry, as the costs of non-herding are self-reinforcing.

In addition to the characteristics, we formulate three general propositions:

General Proposition GP1: Higher costs of non-herding induce higher herding behavior of employees in an organization.

General Proposition GP2: Individual factors of employees, their co-workers and supervisors influence the costs of non-herding for employees.

General Proposition GP3: Institutional factors of organizations affect the costs of non-herding for employees.

3.3. The Influence of Institutional and Individual Factors on the Costs of Non-Herding

3.3.1. Institutional Determinants of the Costs of Non-Herding

We suggest that the costs of non-herding (CONH) can be influenced by various determinants. We pool the determinants of the CONH into three categories in line with the sources of CONH mentioned above: individual and institutional determinants, of which the latter are divided in institutional determinants at the micro and the macro levels. Examples for individual determinants of the CONH include the employee's personality, experience, and knowledge. Institutional determinants influence the employee's CONH due to factors at the micro level, i.e., the employee's relationship and conduct with co-workers and supervisors in daily business, and factors at the macro level such as human resource management, or ownership structure of the organization. Institutions are the fundamentals that shape "the rules of the game" upon which employees act (North 1990, p. 3). These propositions are underpinned with empirical findings from the literature review in Chapter 2. We start with the development of quantitatively testable propositions for the institutional factors of an organization: Hierarchical structure, size, human resource management, risk management and ownership structure.

Proposition P1: The flatter and more decentralized the hierarchical structure of the organization is, the lower are the costs of non-herding for the employees.

This first proposition captures the intuition that, when stronger hierarchical structures are developed in an organization, the CONH are higher. If there are only a few employees or managers at the same hierarchy level, the allocation of power and responsibility is clearly defined, so that there are few incentives to be critical and not to herd. If the organization is dedicated to a more cooperative or participatory style of management, then there are fewer levels of hierarchy, and more employees are at the same level. This provides a platform for discussion among equals and thus may help to reduce the CONH. Stein (2002) captures this idea in his theoretical work on information production in decentralized and hierarchical firms, and Berger et al. (2005) provide relevant empirical evidence. Stein shows that, in hierarchical firms, information can only be processed efficiently when the information can be "hardened" (Stein 2002, p. 1893). "Soft" information is "information that cannot be directly verified by anyone other than the agent who produces it" (Stein 2002, p. 1892). Thus, soft information can only be used efficiently in organizations with flat hierarchies, where soft information can be processed and communicated fast and in a direct manner. In organizations with tall hierarchies, soft information needs to be hardened as it has to be communicated up many hierarchical layers.[19] As we assume that the costs arising from non-herding behavior are similar to the costs of communicating and processing soft information in an organization, we deduce that a flatter and more decentralized organization has lower CONH.

Proposition P2: The larger the organization is, the higher are the costs of non-herding.

This second proposition is close to the first one, as it seems to be natural that smaller organizations have shorter distances of communication and, thus, the employees can convey their private information in a faster and more direct way. Hardening information is less important the smaller an organization is. Furthermore, in smaller organizations, information spreads much faster between non-related teams, and the interactions between employees of extraneous areas is much higher.

Proposition P3: Privately held companies have lower costs of non-herding than publicly listed companies.

We propose that privately held companies lower CONH, because personally liable partners, for example in a private bank, are more interested in carefully aggregating information and therefore reducing the CONH in order to reveal as many critical issues as possible. In contrast, we assume that in a stock corporation with a large free-float,[20] few employees have the incentive or the power to raise critical issues due to the diffusion of responsibility. This proposition is related to the two propositions stated above, as size and hierarchical structure in firms are linked to the ownership structure as well. For example, Guadalupe et al. (2012) explicitly describe a new trend to organize publicly listed firms more similarly to

[19] Stein (2002, p. 1982) exemplifies "soft" information with a loan officer situation: "For example, a loan officer who has worked with a small-company president may come to believe that the president is honest and hardworking – in other words, the classic candidate for an unsecured "character loan." Unfortunately, these attributes cannot be unambiguously documented [i.e. "hardened"] in a report that the loan officer can pass on to his superiors [over many hierarchical layers]."

[20] The free-float of a listed company is defined as the ratio of shares that are readily tradable to outstanding shares. Shares that are not readily tradable are held by controlling or long-term investors, or government entities, for example.

partnerships, which is a common form of ownership for privately held firms. They suggest that this trend stems from the highly competitive industries that these firms operate in. These industries typically use less physical capital but demand more knowledge work. Such an environment seems to be met favorably by partnerships, as they tend to be more decentralized organizations with flatter hierarchies.

Proposition P4: The lower the part of the salary that is performance related, the lower the costs of non-herding.

The third proposition deals with the problem of how to remunerate the employees in such a way that the CONH are reduced. The literature of reputation-based herding suggests that, if a large part of the salary is performance related and performance is measured relative to other employees as a type of benchmarking, behavior that deviates from the benchmark is very costly (Scharfstein and Stein 1990; Zwiebel 1995). Rajan (1994) provides evidence that rational bank managers with short horizons, which are often induced by short-term performance pay, herd in their credit lending policies. Hence, we conclude that incentive pay usually leads to an increase in CONH. Of course, we do *not* aim to say that incentive pay could *not* be developed in a way which could set limits to the CONH or even lower them. However, the last financial crisis provided ample evidence that the performance pay was mostly constructed in inappropriate ways (Rost and Osterloh 2009).

Proposition P5: The more independent and established the risk management in an organization is, the lower are the costs of non-herding.

The last proposition suggests an important institutional factor which we assume influences the CONH. If the risk management inside the organization is firmly established and unifies enough resources, it can systematically scrutinize prevailing opinions or provide employees with a base for expressing critical views. In order to keep the CONH for the risk management team as low as possible, a high independence of the team members from the units which they supervise has to be ensured. The same logic is pursued with a devil's advocate,[21] as he or she plays the institutionalized role of revealing all possible objections against a decision that seems obvious at first sight (see, e.g., Morck 2008). Similarly, Janis (1972, p. 209) proposes that a team leader should assign the role of critical evaluator to every team member in order to

[21] The *Holy Office of the Devil's Advocate* was invented in the middle of the second millenium by the Roman Catholic Church to take a contraposition when members of the clerus intended to canonize someone. The devil's advocate had to argue against the canonization and scrutinize the actions, attitudes and origin of the candidate, so that the church would not canonize a person that might bring disgrace to the organization; also, it prevented powerful individuals from canonizing their family and friends (see, e.g., Stanley 1981).

reduce groupthink. In contrast, the function of a risk manager or a risk management unit is assigned to a specific person or institution inside the organization. Dewatripont and Tirole (1999) provide a model of how agencies, similar to a devil's advocate, can help to improve decision procedures in organizations, and Garicano and Posner (2005) provide anecdotal evidence from the U.S. intelligence system. They refer to the case of a CIA report claiming to have found weapons of mass destruction in Iraq in 2005. The report was built on one single source, which was discredited afterwards. Senior agency managers on the case could cover the failure for a long period as they kept to themselves and did not have to reveal the case to an independent body that could have discovered the deficiency earlier. Garicano and Posner suggest that agency managers could shift regularly. After a shift, the new manager would then function as a devil's advocate (see also Hertzberg et al. 2010 for a similar case with loan officers).

Institutional factors at the micro level

Besides the institutional factors at the macro level, institutional factors at the micro level can also be identified that can help lower the CONH in organizations. Institutions at the micro level usually include the procedure for setting up meetings and selecting committees and teams. We do not focus on these factors and we do not formulate propositions here, as a comprehensive literature already exists focusing on how herding behavior can be minimized at the micro level (see, e.g., Janis 1972; Bantel and Jackson 1989; Morrison 2003; Garicano and Posner 2005; Schulz-Hardt et al. 2006). At this point, we only provide a short overview of the factors that have been raised in the literature cited above. The literature distinguishes the following institutional factors at the micro level: (1) the composition of the committee; (2) the agenda in meetings; (3) the information flow (e.g., sequence in which and time when employees receive the agenda and notes); (4) the formal communication procedure (e.g., order of speaking); (5) informal communication (e.g., possibility of pre-meetings, other informal gatherings of the group); (6) the decision rules (e.g., definition of the meeting leader, or voting procedures); (7) allocation of responsibility (group leader bears full responsibility or responsibility is shared equally among all members).

These factors define the design of a committee and its rules. They have a strong direct influence on the CONH for an individual committee member. Assume that the project leader has the power to select the members of the committee or to set the agenda of the meeting. This enables him or her to raise the costs for a possible critic by electing a higher number of supporters of the project (or even excluding all possible critics). Finally, the leader can place the vote on the most critical project at the very end of the meeting's agenda in order to minimize the time available for a proper discussion and, thus, raising the CONH for an opposing member. The project leader can also heavily regulate the flow of information about the project. In addition, the formal communication rules, such as the order of speaking in a meeting, the decision-making procedure and the allocation of responsibility for the decisions taken are factors at the micro level of an organization. Nevertheless, they can be influenced by decisions and settings at the macro level of the organization.

3.3.2. Individual Determinants of the Costs of Non-Herding

Individual factors of the employee's personality (e.g, neuroticism, extraversion etc.) have an important influence on the CONH, as has been shown empirically in the psychological and sociopsychological literature (Asch 1951; Festinger 1954; Janis 1972; Bowen and Blackmon 2003; Baron 2005). However, we are focusing on individual determinants of the CONH that are non-psychological, such as the age of an employee or the experience on the job (see the theoretical literature in Scharfstein and Stein 1990; Prendergast and Stole 1996; Hirshleifer and Teoh 2003; 2009). There is anecdotal and empirical evidence about various individual factors that increase or lower an individual's CONH (see, e.g., Bowen and Blackmon 2003; Tangirala and Ramanujam 2008). In particular, empirical research employing large datasets on security analysts has examined the interaction of individual characteristics with respect to herding behavior (Graham 1999; Hong et al. 2000; Clement and Tse 2005). We condense the specific findings in the literature into the following propositions on the individual factors that influence employees' CONH.

Proposition P6: The more experienced the employee is, the lower are the employee's costs of non-herding.

The proposition builds on the reputational-herding model which was developed in Scharfstein and Stein (1990) and has been empirically examined for the herding behavior of financial analysts by Hong et al. (2000) and Clement and Tse (2005).[22] They provide evidence that

[22] This stands in contrast to the model by Prendergast and Stole (1996), in which the old employees ("jaded old timers") exhibit more conformity in an organization than the young employees ("impetous youngsters"). The difference in the models stems from Prendergast and Stole's assumption that young managers are incentivized to distinguish themselves with their first decision from the older managers, showing ability by believing in their own signal. Although Scharfstein and Stein's (1990) model is highly similar, it has the decisive difference that the young manager also anticipates that the market (for managers) will also infer his ability from the difference in performance between his call and the call of the other (older) managers. By doing so, the younger manager

more experienced sell-side analysts exhibit less herd behavior than inexperienced analysts. This is mainly due to career concerns; if less experienced analysts have a lower forecast accuracy, they are punished more severely than the more experienced analysts, whose track records grant them a higher reputation. In particular, the less experienced analysts have a higher probability of being dismissed.

Proposition P7: The older the employee is, the lower are the employee's costs of non-herding.

Proposition P7 is related to Proposition P6, as a young employee, by definition, cannot be experienced. However, if analyzed under the ceteris paribus condition holding experience and other factors constant, we predict a negative effect of age on the employees' CONH. The employee has a certain track record and the selection over time provides a natural effect of reducing the CONH. The empirical findings of Tangirala and Ramanujam (2008) on the relation of an employee's motivation to raise objections and his or her age are mixed. So, the factor merits further analysis.

Proposition P8: The longer the tenure of the employee with an organization is, the lower are the employee's costs of non-herding.

The proposition is related to Propositions 6 and 7, as a young employee cannot have a long tenure in a company and cannot be experienced in an area. Still, one can envisage a case where an old employee with a long tenure in a company changes function inside the firm and faces a situation where he or she is old and has a long tenure in a company but not much experience in the new field of activity. We are not aware of an empirical study explicitly analyzing the tenure of an employee in connection with herding behavior.

Proposition P9: The higher the employee's hierarchical position is, the lower are the employee's costs of non-herding.

This proposition is also directly linked to the reputational-herding model in Scharfstein and Stein (1990), as a higher hierarchical position usually comes with a higher reputation. As discussed above, a difference in rank between two employees provokes higher CONH for the employee with an inferior rank.

Proposition P10: The better the employee is educated, the lower are the employee's costs of non-herding.

---

has a strong incentive to go with the herd (i.e. to adapt to the older analysts). The empirical literature supports the latter model more favorably.

We formulate this proposition in an exploratory sense, as we are aware of neither empirical nor theoretical studies that analyze the link between education and herd behavior directly.[23] The rationale in the proposition stems from the idea that education improves the ability to understand complex interdependencies and, hence, to identify possible problems in projects in organizations. Further, a larger knowledge lowers the employee's CONH by increasing his or her reputation. This mechanism empowers the better-educated employees to join the discussion during a decision-making process. In addition, one can argue that better educated employees have better outside options and thus rely less on a specific job than less educated employees

Proposition P11: The wealthier the employee is, the lower are the employee's costs of non-herding.

This proposition, too, is highly exploratory. We are not aware of any research that focuses on herd behavior in organizations in relation to the employees' wealth. We predict a lower CONH for wealthier employees simply due to the fact that they are more independent of the income generated by their job. We test these propositions qualitatively with in-depth interviews with financial market professionals

3.4. Interview Data

To test our theoretical propositions on the CONH, we conducted interviews with practitioners in the financial industry. From June to August 2010, we interviewed twelve practitioners in the greater area of Zurich, Switzerland's financial center. To attain a reasonably broad insight into the various types of financial institutions, we interviewed four employees of large Swiss banks, four employees of mid-sized banks, two employees of a hedge fund company, an employee of a large global reinsurance company, and an employee of a small investment boutique.

In addition, our sample varies with respect to the positions and functions of the employees. One person was a former CEO of one of the two biggest Swiss banks and another was a former chairman of a global reinsurance company (both companies are on the Fortune Global 500 list); four employees worked as security analysts, one as a managing director, one as an

[23] Cipriano and Guarino (2009) show only as a side product in their laboratory experiment with financial market professionals, that the level of education (BA, MA or PhD) does not have a significant influence on the probability of engaging in an informational cascade, i.e. to engage in information-based herding.

investment adviser; two employees worked as chief risk officers, one as CEO of a hedge fund company, and one as chairman of a small investment boutique.

The sample comprises one woman and eleven men. The sample includes companies ranging in size from 30 to over 60,000 employees. The median number of employees is 5,000 and the mean is around 20,000. The revenues of the companies in 2012 range from USD 22.5 million (estimate) to 48 billion.

An outline was used in all the focused interviews[24], but the order in which the questions were posed varied depending on the course of the conversation. The shortest interview lasted about one hour, the longest about one hour and a half. The interviews were analyzed by applying scaled and structured content analysis (Mayring 2003; Glaeser and Laudel 2006). Appendix A provides an extensive overview of the interview procedure, a discussion of possible biases in interviews, and a detailed description of the evaluation of the interview data.

3.5. Interview Results

3.5.1. General Propositions and Characteristics of the Costs of Non-Herding

In general, all participants confirmed that problems of repressed voice and herding behavior occur in the decision-making processes of organizations. They agreed that there were individual costs involved which lead to silence on the part of employees and to herding behavior. All respondents stated that they had experienced a wide range of costs for non-herding during their careers, starting with delicate psychological pressure to conform and culminating in dismissals. Several respondents made such remarks as the "bears got mobbed,"[25] or "everybody knew the critics and sooner or later they were no longer invited to the important workshops" or "from then on your days are numbered."[26] Responses like these reveal additional types of costs such as a deterioration of career opportunities and reduced bonuses due to criticism and conflicts.

Hence, we conclude that the respondents confirm our first two general propositions, namely that *GP1 higher costs of non-herding induce a higher herding behavior of employees in an*

[24] The term was coined by Merton and Kendall (1946); the procedure is explained in Kvale (1996).
[25] The term "bears" refers to people who forecast an economic downturn; see, for example, *The Economist* (2009).
[26] All the interviews are conducted in German. The excerpts presented here were translated by the authors.

organization; and that *GP2 individual factors of employees (and their co-workers and supervisors) influence the costs of non-herding for the employees*.

All of the twelve interviewees mentioned various institutional factors when asked broadly about how to reduce CONH. A managing director of a major Swiss bank pointed to the hierarchical structure of an organization, without being explicitly asked: “For example, one of our competitors in the market for whom I worked, had a much steeper hierarchy and – I do not want to say fear, this would be exaggerated – but you feel much more uncomfortable about raising your voice and about expressing your contrary opinion.” An analyst working at a mid-sized bank stated: “Well, we do not have such high pressure as in the larger banks, and we do not exert pressure on ourselves – also because we are not remunerated in a highly variable way, or at least to a much smaller extent”.

We conclude that our third general proposition is also supported by the respondents – *GP3 Institutional factors of organizations affect the costs of non-herding for employees.*

Regarding the characteristics of the CONH (Characteristics C1-C6), all managers responded that they perceived the costs to be self-reinforcing over time C6. The CEO of a hedge fund company commented: “It gets more and more difficult, especially if you did not say anything at the beginning [of the project]; well, this lies in the nature of things.” Other characteristics were also addressed, such as the temporal asymmetry between short-term costs and long-term benefits C3, by five of the twelve respondents; and the asymmetry in the probabilities between highly probable costs and very uncertain benefits C4, by six of the twelve respondents. A security analyst explained that sometimes he does not speak up to avoid forcing others to stay longer in the meeting, as he knows that this would provoke disputes and annoyance. Yet, he would remain silent although he does not fear any ultimate consequences and intrinsically admits that speaking up would be more beneficial for the aim of the meeting. He keeps quiet “just as a matter of kindness”, and because he likes a peaceful environment. This supports the first Characteristic C1 stating that the CONH are mostly obvious before the decision is taken; all respondents stated that they perceived such costs ex-ante. Half of the respondents brought up the issue of ex-post justification when they raised their voices C2. Finally, all respondents mentioned the fifth Characteristic C5 that non-herding behavior is costly for the specific employee in various forms. As a former CEO of a large Swiss bank put it: “I know an analyst who has lost his job because he revealed unpleasant things – I mean this is an extreme reaction but these are the private costs to be borne by such an employee alone.”

3.5.2. Institutional Determinants of the Costs of Non-Herding

Concerning our propositions of institutional factors at the macro level, the respondents exhibited a more heterogeneous picture. The question regarding the hierarchical structure of the organization – *Proposition P1: The flatter and more decentralized the hierarchical structure of the organization is, the lower are the costs of non-herding for the employees* – was affirmed by nine of the twelve respondents. “In general, I think that younger firms have an advantage, as they have less established structures and hierarchies”, a managing director of a major Swiss bank answered. We conclude that a flatter and more decentralized hierarchical structure lowers the CONH in organizations.

Only four interviewees agreed that the size of the company had an influence on the CONH; the remaining eight respondents disagreed. One respondent claimed: “This is not a matter of size, normally even in big banks the daily work environment does not differ between a big bank and an institute of small or medium size.” Hence, we do not find evidence that the size of the organization influences the CONH. The respondents denied Proposition P2*: The larger the organization is, the higher are the costs of non-herding.*

When asked if the ownership structure can influence the CONH, seven of the respondents confirmed the proposition and five of them disagreed with it. Although a minority did not agree, we weakly confirm that the majority of the respondents supported Proposition P3 – *privately held companies have lower CONH than publicly listed companies.* An investment adviser explains:

> “My employer [a major Swiss bank] provides a nice example. […] our shareholders were atomized and finally, nobody is responsible for anything. So you drift into anonymity. If you work with a small private bank, you know that the partners do not want to invest in U.S. mortgages and other mortgage backed securities. They don’t have to engage in trends and can resist going with the herd.”

On the issue of an organization’s wage policy, nine of the twelve respondents were convinced that a lower performance-related component of the salary decreased the CONH. One interviewee explained his view by stating that contributing private information was a public good and the existence of large personal incentives sabotaged the provision of these kinds of goods. In contrast, one of the three respondents who went against the proposition noted that only the existence of performance pay kept him attentive as he could have lost a great deal of money when not providing all of his private information. However, we conclude that Proposition P4 – *the lower the part of the salary that is performance related, the lower the CONH* – was supported.

Ten out of the twelve respondents confirmed Proposition P5 – the more independent and established the risk management in an organization is, the lower is the herding behavior in the organization. A risk officer of a mid-sized Swiss bank noted: “In our company, the CFO is also the Chief Risk Officer (CRO), and he sits on the executive board and also reports directly to the board of directors; I think this is essential.” Or, as the former chairman of a global reinsurance company put it: “From my point of view, a CRO belongs to the executive board as an independent unit.” An analyst of a mid-sized bank concluded: “If the CRO is positioned too close to the business, you should not dare to hope that he can influence anything.” Hence, Proposition P5 is strongly supported.

Institutional determinants of the costs of non-herding at the micro level

Although we did not formulate institutional propositions at the micro level, we briefly report our findings in this area to provide a complete picture of determinants influencing the CONH. Ten of the twelve interviewees spontaneously brought up several points at the micro level when asked which institutional factors influenced the CONH in meetings. All of the seven factors mentioned at the micro level were addressed by at least one respondent. Surprisingly, although the answers took many forms, the respondents stressed the importance of informal communication (5) before or after the meeting (whenever a decision has not yet been made) in reducing or avoiding CONH. In addition, the formal communication procedure (4) was the second most mentioned institutional factor at the micro level.

Without being asked, the CEO of a small hedge fund company explained their rules on how to proceed in investment committee meetings: They try to rule out every imbalance by randomizing the order in which the topics are discussed or who starts the discussion. In addition, all of the members of their investment committee have veto power. Interestingly, over the period of four years, a veto has been cast only once. The respondent stated that the mere existence of the veto leads them to consider each member’s opinion and to empower each of the members to raise his or her voice if there is a disagreement with a proposed decision.

The majority of the respondents were aware of the importance of institutional factors at the micro level. However, in their companies, rules about the procedure of meetings or other decision-making boards were ill defined or not adequately implemented (except in the hedge fund firm). A managing director of a major Swiss bank stated sharply: "It is official that the firm wants this [employees to be able to speak up]; however, it depends very much on the implementation, on the kind of boss you have and on the person who leads a meeting." The former CEO of a large Swiss bank mentioned that during his time in office every member of the executive board had to sign a declaration that obliged him to dissent with other members whenever he had a different opinion. Although this was not a credible gambit, the CEO wanted to make clear that dissent was important in finding the right solutions and taking proper decisions. We record these responses and infer that various institutional factors at the micro level influence an employee's CONH.

3.5.3. Individual Determinants of the Costs of Non-Herding

Eleven of the twelve respondents agreed with the first proposition on the employees' experience, P6 – *The more experienced the employee is, the lower are the employee's costs of non-herding.* This stands in contrast to Proposition P7 – *the older the employee is, the lower are the employee's costs of non-herding* – to which the respondents answered in a highly ambivalent way; six respondents supported the idea that older employees have lower CONH while five did not, and one interviewee even refused to give a distinct answer. The answers on Propositions P6 and P7 reveal the asymmetric dependency between age and experience. Hong et al. (2000) found empirically that experienced financial analysts were more willing to resist the prevailing opinion. They faced a lower probability of being dismissed after having lower accuracy; hence, they bore lower CONH. However, Hong et al. (2000) do not clearly distinguish between age and experience in their argumentation, which seems to be a shortcoming with regard to our results. The interviews revealed that it was the experience in connection with the high reputation that produced lower CONH for older employees, not age itself. This statement is in line with the literature on reputation-based herding (Scharfstein and Stein 1990) and on voice and silence in organizations (Morrison and Milliken 2000).

As in the case of age, the responses to Proposition P8 – *the longer the tenure of the employee with an organization is, the lower are the employee's costs of non-herding* – also revealed a heterogeneous picture with a slight tendency to support the proposition. Seven respondents agreed, while five respondents did not. Two of these five respondents stated that the effect of higher tenure led in the opposite direction: The longer the tenure, the higher the CONH for the employee. A risk officer of a mid-sized Swiss bank stated: "Well, I see both ways; either the longer tenure provides you with a higher identification and loyalty to your employer [i.e. non-herding], or you get tired and do not want to "hurt" anyone [i.e. herding]." The seven respondents who supported the proposition argued that the costs of speaking up were lower for an employee who had been with the organization for a long time, as the employee must have survived many years and thus must enjoy a particular status and a high reputation within the organization. One manager of a major Swiss bank stated: "There were issues, a year ago, which I avoided commenting on, because I knew perfectly well that I was the newcomer and all the other co-workers knew much more than I did. Now, after one year, I have a much higher impact on the decision-making processes and dare to speak up." This example might point to a positive but diminishing marginal effect between tenure and the CONH in an organization. In the beginning, the growing tenure lowers the employees' costs of speaking up drastically. Later, the positive impact of tenure on the CONH decreases.

The next Proposition, P9 – *the higher the employee's hierarchical position is, the lower are the employee's costs of non-herding* – was fully supported by all the respondents. A security analyst remarked:

> "Yes, a higher hierarchical position lowers the employee's CONH naturally. But this reveals that informal communication is so important that people should know each other across the company and across hierarchies, so that people of a lower rank can talk to more senior people and point to problems which the high ranked executives do not know or see."

The proposition on education was not supported by the respondents. Only four out of twelve managers agreed with Proposition P10 – *the better the employee is educated, the lower are the employee's costs of non-herding.*

Three remained undetermined and five interviewees denied the relationship between an employee's education and his or her CONH. As we cannot draw on literature relating the CONH of an individual to education, we need to reconsider this argument.

In contrast, our final Proposition P11 – *the wealthier the employee is, the lower are the employee's costs of non-herding* – was fully supported. All respondents pointed to the higher independence of a wealthy employee in contrast to a poor employee. The fear of not being able to deal with and financially survive a dismissal creates high CONH for the employee. A chief risk officer pointed out: "In our industry, this is absolutely true; also because wealthy people have a very critical point of view on financial markets as they have to manage their own wealth." The only critical remark was mentioned by an analyst: "One might argue that sometimes, rich people tend also to increase their fixed costs strongly and thus face even higher CONH and fear of losing their job. But in general, higher wealth should lead to lower CONH for the employee."

Finally, when asked if they were aware of companies with a human resource practice, which sought to hire people according to specific individual factors in order to maintain critical thinking in an organization, only two respondents agreed. One of them, a CEO of a hedge fund company, even introduced such a policy in his firm. The second respondent, a former CEO of a major Swiss bank, remarked that "You have to do that, you have to select independent thinkers. Everything else is boring and counterproductive." However, we cannot conclude that human resource policies that aim at selecting independently thinking employees are an important institutional determinant, as the large majority of our respondents have no experience with it. Yet, it might be fruitful for directors in organizations to develop a human resource strategy which sought to prevent herding behavior among their employees.

3.6. Conclusion

Employees' herding behavior has to be overcome for an efficient decision-making in organizations. Whenever employees exhibit herding behavior in an organization, they hide their private information. This loss of information deteriorates the organization's decision-making process, as the decisions taken rest upon a less complete information basis (Argyris and Schön 1978; Scharfstein and Stein 1990; Miliken et al. 2003; Hirshleifer and Teoh 2009). The question is: Why do employees not contribute their private information and instead hide in the herd?

In this chapter, we argue that costs arise for employees if they do not follow the herd. Using in-depth interviews with high-ranked executives and analysts in the financial sector, we explored various facets of the costs of non-herding (CONH). The respondents stated that they experienced a wide range of costs during their careers, starting with delicate psychological pressure to conform and culminating in a call to quit a job or even in dismissal. Four important characteristics leading to asymmetries of the costs of non-herding are identified and supported by the respondents. Firstly, the costs of non-herding bear the asymmetric dynamic of a self-reinforcing mechanism. Each employee that joins the herd increases the costs on all other employees that are still staying outside the herd. Secondly and thirdly, the costs of non-herding possess a twofold asymmetrical relationship to possible benefits: the costs are imminent and forseeable, whereas possible benefits lie in the future and are uncertain. Finally, the costs of non-herding bear a characteristic of a public good. The costs have to be borne individually while the benefits have to be shared with other members of the organization. While the first aspect of the asymmetric dynamic is a characteristic of every herding behavior situation, the latter three asymmetries are features typical of organizational contexts in which hierarchical and economic dependence and a common goal exacerbate the impact of the costs. The interviewees concluded that the combination of these asymmetries makes the costs of non-herding so difficult to predict and the research for means to master them important.

Further, the analysis of the interviews revealed that institutional factors of organizations substantially influence the employees' CONH in an organization. The respondents confirm in particular that a flatter hierarchy, a private ownership structure, and less performance pay in organizations help to reduce the employee's CONH. In addition, executives argue that a human resource policy which aims to select employees with individual factors that lower their CONH might succeed in improving an organization's information aggregation. The interviewees indicated various individual determinants that lower the personal CONH for an employee. Foremost, all the respondents argued that a higher experience would diminish the CONH, while age and tenure, ceteris paribus, have an ambivalent influence. Interestingly, the respondents declined to acknowledge a direct influence of employees' education on their CONH; however, they confirmed that the employees' wealth has a negative influence on the CONH, leading to less herding behavior in an organization.

We conclude that the employee's CONH are key in overcoming herd behavior in organizations. Lower CONH will provide a larger amount of information to base the organization's decisions on. When the CONH are high, most of this information rests unrevealed in the employees' minds. Most importantly, according to the responses of the executives and analysts in our sample, institutional factors of organizations do influence these CONH and can lower them. These insights should be applied in practice.

Furthermore, the CONH lead to pressure on the individual employee to remain silent. Morrison and Milliken (2000) list several implications of the costs of voicing objections (i.e.

CONH) and the resulting pressure on employees. First, employees have feelings of not being valued. Second, employees perceive a shortage of self-control. Third, employees experience cognitive dissonance. These impacts of the CONH also deteriorate the organization's performance in an indirect way by reducing an employee's motivation at work (Parker 1993). Happiness research in economics has provided evidence that an employee's perception of higher self-control and autonomy fosters an employee's job satisfaction (Benz and Frey 2008; Frey 2008). In order to employ more information for the decision processes and to bring forth more satisfied and motivated employees, managers ought to reduce the CONH faced by their employees. Lowering the CONH would indirectly foster employees' motivation by simultaneously enhancing the organization's efficiency in the decision-making processes.

Other studies have analyzed how managers can support voice more effectively and reduce herding behavior in organizations by changing their beliefs, practices, and fear of negative feedback (Milliken et al. 2003; Morrison and Milliken 2003). Our study provides insights on the institutional and individual determinants of employees' CONH that managers should focus on. Managers that aim at increasing a firm's aggregation of private employee information could favorably select employees according to the individual determinants of the CONH presented here and might adjust the organization towards the stated institutional determinants. Future research should explore the relationship between the various institutional factors and the different individual determinants in order to enhance the informational basis of the decision-making process in organizations.

4. How Institutional Factors of Firms Influence Employees' Herding Behavior: The Case of Financial Analysts[27]

4.1. Introduction

The aggregation of private information possessed by individuals is among the greatest challenges to be resolved in the social sciences, according to the scientific journal *Nature* (Giles 2011). The aggregation of independent opinions and forecasts generally leads to a highly accurate estimate of the actual outcome, which is often called "the wisdom of crowds" (Galton 1907; Surowiecki 2005). However, social influence can have detrimental effects on the aggregation of individuals' private information (e.g., Lorenz et al. 2011). The pressure to conform can create a herding behavior that destroys the positive effects of information aggregation. Among other sociopsychological factors, reputational concerns are an important force pushing individuals to follow the herd.[28] When individual forecasters know that their ability is judged by comparing their forecasts to the aggregated forecasts of others, herding can become a dominant strategy for every forecaster in the market (Scharfstein and Stein 1990). Thus, individuals have incentives to bias the public expression of their private information towards the opinion prevailing among the majority when they know that their ability or reputation in the market is assessed in such a relative way.

One field in which the aggregation of information is crucial is the financial market. A distortion of information aggregation destroys the basis for efficient asset and resource allocation, and, hence, shapes not only the financial sector but also all other sectors linked to capital markets. As security analysts[29] play a key role in providing information to capital

[27] Parts of this chapter are based on a paper written jointly with Peter Fiechter (Cueni and Fiechter 2013).
[28] Please refer to Chapter 2 for an extensive review of sociopsychological factors that can drive herding behavior of individuals.
[29] The term security analyst indicates an analyst that is covering stocks and other securities in financial markets.

markets, they offer an ideal setting in which to analyze herding behavior and its determinants (e.g., Chang et al. 2006).

Analysts' herding behavior results in a loss of information, increases information asymmetry, and ultimately leads to misallocation of scarce resources (e.g., Trueman 1994). To avoid this waste of resources, it is important to understand the drivers of analysts' herding behavior. Prior literature on herding behavior has focused on individual analyst characteristics. For instance, Hong et al. (2000) and Clement and Tse (2005) have shown that individual factors such as job experience and self-assessed ability influence the herding behavior of analysts. However, the literature on analysts' herding behavior is largely silent about the influence of the institutional environment. The size of the broker (i.e. number of analysts employed) as a proxy for reputation or available resources is the only institutional factor appearing in previous research on analysts' herding behavior.[30] As research from other fields in economics generally finds a strong impact of various institutional factors on human behavior and well-being (e.g., Veblen 1899; Williamson 1975; Frey and Stutzer 2002b), we expect that institutional factors influence analysts' herding behavior beyond the individual analyst characteristics. In addition, recent anecdotal evidence also suggests that the institutional environment affects analysts' forecasting behavior. For example, *Bloomberg Markets Magazine* subtitled one of its cover stories "Fed up with big bank's conflicts, analysts are starting their own firms", explaining that analysts "[…] grew fed up with a culture that prized irrational exuberance over sober analysis" (Robinson 2009, p. 46).

This study examines the influence of institutional factors on the herding behavior of security analysts. First, to achieve this goal, a general framework combining individual and institutional factors is presented. The framework analyzes the costs of non-herding (CONH) that influence the forecasting behavior of analysts (and of other information-providing agents in general). Low CONH indicates that the costs of not following the herd are low for a particular analyst, and that the analyst is thus less prone to herding behavior than other analysts facing higher CONH. Second, based on the framework, we introduce and examine the influence of a new institutional factor, *ownership*, on the herding behavior of security analysts. Specifically, we investigate, firstly, whether analysts working at publicly listed

[30] We do acknowledge that the literature on analysts' forecast *errors* addresses institutional factors such as brokers' underwriting and investment banking activities (O'Brien et al. 2005; Cowen et al. 2006; Ljungqvist et al. 2006; Malmendier and Shanthikumar 2007, Barber et al. 2007), the geographical distance between the analyst and the covered company (Malloy 2005), and financial incentives in relation to analysts' accuracy (Groysberg et al. 2011).

brokerage houses are more likely to herd than analysts from privately held brokers, and secondly, if this effect is different for analysts covering their own sector (i.e., financial sector stocks) compared to analysts covering other sectors.

First, we predict that analysts from public brokers are less independent and, thus, more prone to herding behavior than analysts from private brokers, because analysts from public brokers face higher costs when processing soft information, i.e. information that "cannot be directly verified by anyone other than the agent who produces it" (Stein 2002, p. 1892). Soft information is crucial in establishing an independent opinion. The processing of soft information is more difficult in more centralized and regulated environments with taller hierarchies which hold the scope and size of the brokerage fixed (Stein 2002; Berger et al. 2005). Therefore, we expect the CONH to be lower for analysts from private brokers than for analysts from public brokers, as the former companies generally share flatter hierarchies and operate in a more decentralized manner than publicly listed brokers. Lower costs are expected to lead to less herding behavior among analysts.

Second, we expect that the difference in CONH between analysts from publicly listed and privately held brokers is substantially larger for analysts covering the financial sector. These analysts could face higher costs because their forecasts influence not only the valuation of the covered firms in the financial sector but also the valuation of their own employer, as their employer is listed in the same sector. Forecasts of financial sector analysts[31] thus exert an externality on their employer's stock valuation, thereby affecting the bonuses and career prospects of their own executives. Hence, we predict that the difference in herding behavior between analysts from publicly listed brokers and analysts from privately held brokers is larger among analysts covering the financial sector than among analysts covering other sectors.

To test our main research question, we investigated analyst forecast revisions[32] from 1999 to 2008. We selected this time period because the last financial crisis has provided many anecdotes of security analysts who did not herd and were bullied or even dismissed as a result (see, e.g., *The Economist* 2009; Robinson 2009). We constructed the herding measure following Gleason and Lee (2003) and Clement and Tse (2005). The empirical proxy

[31] The term *financial sector analysts* stands for analysts that are covering stocks of companies from the financial sector, like banks and other financial services firms.

[32] During a financial period (e.g., a financial year), an analyst usually revises his or her forecast on a certain stock several times in order to adjust his or her prediction according to new information about the listed company. We focus on analysts' earnings per share (EPS) forecasts of listed U.S. companies.

classifies forecast revisions of earnings as *bold* if the analyst's forecast revision deviates from both the prior forecast and the prerevision consensus, and as *herding* if the forecast revision is issued between the analyst's prior forecast and the prerevision consensus. The prerevision consensus is the averaged forecast of all other analysts covering the same stock as the specific analyst, before the specific analyst revises his forecast.[33]

To test the effect of the broker's ownership, we hand-collected data on whether U.S. brokers are publicly listed or privately held and matched them with data on analysts' earnings forecasts. We enriched this unique dataset with data about the brokers' underwriting business[34] and their informational environment and included a full set of controls for the individual analyst characteristics. This allowed us to regress the herding measure on the ownership variable by controlling for differences in informational resources and reputation across public and private brokers, other institutional brokerage characteristics, and a full set of variables capturing individual analyst characteristics.

First and foremost, we find a substantial effect of the brokers' ownership on the herding behavior of their analysts. Second, the findings reveal a tradeoff between the brokers' size and the brokers' ownership. At small brokers, the lack of resources and the lower reputation seems to make their analysts more prone to herding behavior; for privately held brokerages, the effect is more accentuated. However, this overall effect holds until a threshold of about 40 analysts. Analysts working at a privately held brokerage that employs more than 40 analysts show a significantly lower probability of issuing herding forecasts of 2.6%. These findings reveal that analysts from private brokers face lower CONH on average, resulting in a lower propensity to herd, but also that this ownership effect does not become evident until a certain level of available resources is reached.

Third, we can show that the ownership effect is driven by the group of analysts covering the sector they are working in, namely the analysts covering financial sector stocks. Again, the probability of issuing a herding forecast decreases with the broker's size for analysts at both types of brokers, but for analysts at privately held brokers the decrease is much stronger. Hence, the difference in the probability of issuing a herding forecast between financial sector analysts working at private brokers and those at public brokers becomes significant at the level of broker size of 24 analysts. At this level, the probability of issuing a herding forecast is

[33] Please also refer to the data section (4.3) and Figure 1 for a more extensive explanation of the herding measure employed in this chapter.

[34] Many brokerage houses are engaged in underwriting business that mostly includes services for initial public offerings and seasoned equity offerings of firms (debt issuance and equity issuance).

3.0% lower for a financial sector analyst at a private broker than for the same analyst at a public broker. At a higher level of broker size of 60 analysts, the probability of herding for financial sector analysts from a private broker is 6.5% lower than for the same analysts at a publicly listed broker. To reach the same effect, for example, an analyst with 1 year of professional experience would have to gain more than 26 years of additional experience. These differences in the probability of herding account for an economically meaningful impact of the broker's ownership on analysts' CONH and, hence, on their herding behavior.

Fourth, we can confirm the findings by Clement and Tse (2005) that bold forecasts are more accurate than herding forecasts, and not simply the result of overconfidence on the part of poorly informed financial sector analysts. Finally, the effect is robust to a battery of additional checks, all confirming our results.

The insights we gain from our study help to understand how institutional factors in brokerage houses influence herding behavior. The institutional factor of ownership addresses not only a different channel from the institutional determinants already used in the literature, a broker's reputation or available resources, but also a different channel from a broker's investment banking business, as used in the forecast error literature. Our findings show that the environment at publicly listed brokerage houses constrains financial sector analysts, increasing their CONH and hindering them from revealing their private information. This impedes efficient information aggregation and biases the investors' view of the future performance of the financial sector. As we show that a more independent analysis is provided by financial sector analysts working with privately held brokers, our findings inform investors and regulators about the factors that lower the analysts' CONH. Thus, a higher attention to forecasts from privately held brokers might support a more genuine and independent forecasting of the future performance of financial sector stocks, enabling a more efficient asset and resource allocation in financial markets.

The chapter proceeds as follows. In Section 4.2, we review the related literature and introduce ownership as an additional institutional determinant of analysts' herding behavior. Section 4.3 describes our data. Section 4.4 explains the methods and predictions. Our results and the according robustness checks are presented in Section 4.5. Section 4.6 concludes.

4.2. Literature, Framework and New Institutional Determinant of Analysts' Herding Behavior

4.2.1. Individual Analyst Characteristics and Analysts' Herding Behavior

The job of sell-side analysts is to provide information about the stocks they cover. Analysts' forecasts appear to reveal valuable private information for investors (e.g., Womack 1996; Kelly and Ljungqvist 2007). For investors, distinguishing between valuable and valueless information is difficult, as analysts' forecasts can be biased due to various distorting incentives.[35] A recent example was the discussion about the analyst Scott Devitt, who covered the IPO[36] of Facebook in 2012 for lead underwriter Morgan Stanley. When Devitt lowered his earnings estimates for Facebook, there was uproar in the industry. The press called Devitt the "sober man in world of hype"[37], as no one expected that the analyst of the lead underwriter would reveal pessimistic information days before the biggest IPO in history – thus clearly confirming that investors perceive analysts as biased (Toonkel 2012).

In general, analysts' performance can be measured ex-post by analyzing the forecast errors, which provides analysts with an incentive to issue accurate forecasts. This incentive matches the investors' incentive to receive valuable information as the basis for decisions about when and how to invest. However, biased forecasts cannot be detected ex ante, and it also remains unclear ex post how much information was available to the analysts at the time the forecasts were made. Because of this lack of knowledge, the ability of analysts and their peers is not usually assessed in absolute error terms but in a relative measure, that is, by benchmarking.[38]

In their seminal paper, Scharfstein and Stein (1990) show that such performance benchmarking leads to herding behavior among analysts.[39] In Scharfstein and Stein's model,

---

[35] For a detailed description about the job and environment the sell-side analyst works in, refer, e.g., to Michaely and Womack (2005).

[36] Facebook conducted its initial public offering (IPO) on May 18, 2012 at the Nasdaq stock exchange.

[37] Devitt was called "sober man in world of hype" by several news agencies, see, e.g., Jessica Toonkel's article "Morgan Stanley's Facebook analyst: Sober man in world of hype" for Reuters (May 31, 2012), retrieved from http://www.reuters.com/article/2012/05/31/us-facebook-morganstanley-devitt-idUSBRE84U04U20120531, on September 5, 2013.

[38] A substantial number of studies support the view that forecast accuracy is usually not that important for sell-side analysts as long as their forecasts do not reach extreme errors, while how "extreme" is defined is usually a relative concept (see, e.g., Clement and Tse 2003; Groysberg et al. 2011). Again, this notion supports the importance of analyzing herd behavior among sell-side analysts and of identifying the characteristics that empower analysts to reveal their full private information.

[39] Herding behavior has also been studied in other contexts, such as informational cascades (see, e.g., Bikhchandanani et al. 1992; Banerjee 1992; Welch 1992) or availability cascades (Kuran and Sunstein 1999); for

analysts always herd, as it is "better to fail conventionally than to succeed unconventionally" in Keynes's words (1936, p. 158). In contrast, in Zwiebel's (1995) extended model it is always the best strategy to succeed. However, both models predict that, for the majority of analysts, who are of mediocre ability, hiding in the herd is the dominant strategy due to ex-ante uncertainty about the future state of the world.

Hong et al. (2000) empirically test the reputation-based herding behavior models of Scharfstein and Stein (1990) and Zwiebel (1995). They proxy the analysts' reputation with the reputation of the brokerage house they work for by counting the number of analysts that are employed at the same house. The larger the number of analysts, the higher is the reputation of the brokerage and, thus, the higher is the individual analyst's reputation. They categorize the analyst's behavior by measuring whether the analyst's forecast moves away from the consensus (i.e. bold behavior) or towards the consensus (i.e. herding behavior). They show that, *after* controlling for forecast accuracy, inexperienced analysts provide less bold forecasts and that they are more likely to be dismissed when issuing a bold forecast. Thus, inexperienced analysts face stronger incentives to herd than experienced analysts, because the inexperienced face higher costs if they neither herd nor make accurate enough forecasts. In addition, they provide evidence that being bold and inaccurate leads to bad career outcomes, while being bold and accurate does not significantly boost analysts' career prospects. Hence, reputational concerns drive inexperienced analysts to herd because they cannot draw on their established reputation in the market.

Clement and Tse (2005) confirm the association between analysts' reputations and their herding behavior and expand the finding by including additional individual analyst characteristics. These are based on the model by Trueman (1994) who proposes that analysts' self-assessed ability drives herding behavior. Clement and Tse (2005) find that the probability that analysts reveal their full private information (i.e., that they issue a non-herding forecast) increases with the experience of the analysts, their last years' forecast accuracy (i.e., their self-assessed ability), and the frequency with which they issue forecasts; it decreases with both the number of industries or firms covered and the days elapsed since the prior forecast of another analyst.[40]

---

a general discussion of the context of social learning and herding behavior in financial markets see, e.g, Hirshleifer and Teoh (2003; 2009). Please refer also to Chapter 2 for an extensive discussion of studies on herding behavior.

[40] Jegadeesh and Kim (2010) support the positive correlation of frequency on analysts' bold behavior of their stock recommendations. Also Graham (1999) studies the association of herding behavior and individual analyst

In contrast to Clement and Tse, Bernhardt et al. (2006) find that analysts in general exhibit anti-herding behavior. However, as our study focuses on forecast revisions, it is important to note that there is no discrepancy between both studies with regard to forecast revisions. Bernhardt et al (2006) and Clement and Tse (2005) both show that revisions moving further away from the consensus are more likely to accurately reveal new private information to the market than revisions that converge with the consensus. Overall, the association between the individual analyst characteristics and the analysts' herding behavior is well grounded in the theoretical and empirical literature.

4.2.2. Institutional Brokerage Characteristics and Herding Behavior

The literature on herding behavior of analysts employs only one institutional characteristic, that is, the broker's size measured by the number of analysts employed.[41] Hong et al. (2000) includes the size of the brokerage as a proxy for whether the analysts experienced a favorable job separation. They imply that joining large brokers increases analysts' reputations and prestige (see also Hong and Kubik 2003).[42] Aside from the use of broker size as a measure of institutional reputation, Clement (1999) explains his finding that analysts working at a larger brokerage exhibit a higher forecast accuracy with the larger resources available to them when making forecasts. Clement and Tse (2005) show that the more analysts working in a brokerage house, the lower is the probability that the individual analyst issues a herding forecast. In contrast, Jegadeesh and Kim (2010) find evidence that analysts working with larger brokerage houses exhibit more herding behavior in their stock recommendations[43].

---

characteristics in a model similar to Scharfstein and Stein (1990). Welch (2000) tests analysts' herding behavior empirically but not explicitly reputational-herding in the sense of Scharfstein and Stein (1990).

[41] We are aware only of the study by Chen and Jiang (2006) that partly includes investment banking business controls and analyzes a topic similar to herding behavior. However, they specifically study the analysts' over- and underweighting of their private and of public information when making forecasts. Furthermore, they do not include controls on the level of the brokers' underwriting business but at the level of the individual stock the analyst covers.

[42] Not within the context of analyst herding behavior, but in a sociological study about conformism, Phillips and Zuckerman (2001) show that analysts exhibit a substantial degree of conformity when issuing sell ratings, but also that the analysts' reputations have a strong influence on conformism. Analysts with very low or very high reputations have a higher probability of issuing a sell rating, thus conforming less. In a study analyzing 216 cases of alleged corporate fraud, Dyck et al. (2010) provide evidence that highly reputed analysts tend to blow the whistle more often. They link this finding to the enhanced information resources of large, prestigious brokerage houses. This indicates the importance of the information resources available to analysts when they want to make a credible statement against the opinion prevailing in the market.

[43] Stock recommendations are the second type of forecasts that security analysts are releasing, besides the EPS forecasts. Normally, the recommendations indicate if the stock should be bought or sold, and are usually labeld with "sell", "hold", or "buy" (or additionally "strong sell" and "strong buy"). In the literature, stock recommendations are seen as investment advice for retail or small-investors, while the EPS forecasts are seen as advice for institutional investors and investment professionals (see, e.g., Malmendier and Shanthikumar 2009).

Importantly, the studies use different herding measures, earnings estimates per share versus stock recommendations. Because analysts use stock recommendations and earnings estimates as different signals to the market (Malmendier and Shanthikumar 2007), the herding behavior might also differ. Therefore, the results reported by Clement and Tse (2005) and by Jegadeesh and Kim (2010) should be compared with caution.

In contrast to the literature on analysts' herding behavior, the literature on analysts' forecast *errors* provides several studies emphasizing the importance of investment banking relationships of brokerage houses (Lin and McNichols 1998; Daniel et al. 2002; Michaely and Womack 2005). The studies suggest that analysts' forecasts can be of low accuracy when the broker incentivizes the analyst to boost its underwriting business. The analyst could, for example, issue favorable forecasts for the stocks of underwriter clients or hostile forecasts for the stocks of clients' competitors. The direct impact of investment banking business on affiliated sell-side analysts is discussed controversially in the analyst literature (see, e.g., Jackson 2005; O'Brien et al. 2005; Cowen et al. 2006; Ljungqvist et al. 2006; Malmendier and Shanthikumar 2007; Barber et al. 2007; Mehran and Stulz 2007). O'Brien et al (2005), for example, show that underwriting business provides incentives for affiliated sell-side analysts to distort their forecasting behavior. In contrast, Cowen et al. (2006), provide evidence that the incentive of brokers to maintain their reputation can reduce the forecast errors of sell-side analysts, and hence, affiliated analysts issue less biased forecasts than unaffiliated analysts.

However, the literature of investors' reactions to analysts' forecasts reveals a largely clear view, showing that investors considerably discount the forecasts of analysts working for brokerage houses with a substantial underwriting business (see, e.g., Lin and McNichols 1998; Michaely and Womack 1999; Malloy 2005; Barber et al. 2007; Agrawal and Chen 2008). This finding supports the notion that the investment banking business negatively affects analysts' forecasting behavior or, at least, that it is perceived as such by investors. Thus, studies designed to analyze analysts' forecasting behavior should control for the investment banking business done by the brokerage houses at which the analysts are employed.

A last institutional factor that is discussed in the literature on analysts' forecast errors is the influence of analysts' pay. It might appear counterintuitive, but the existing evidence suggests that analysts' pay is not strongly linked to forecast accuracy (Groysberg et al. 2011).[44] In contrast, Mikhail et al. (1999) and Hong and Kubik (2003) provide evidence that analysts who issue more accurate forecasts experience more favorable job separations. But, on the same tier of brokerage houses, analyst pay is generally not related to forecast accuracy (Groysberg et al. 2011). Only analysts who issue highly inaccurate bold forecasts have a much higher probability of experiencing an unfavorable job separation. Thus, analysts from a certain tier of brokerage houses are party to a Mirrlees contract, in which forecast accuracy within a normal range is not crucial for pay. This provides incentives to go with the herd, since the herd defines the *normal* range of accuracy (see, e.g., Bolton and Dewatripont 2005, p. 304). In contrast, Groysberg et al. (2011) show that, in a large investment bank, the analysts' ability to promote their employers' investment banking business (in particular, underwriting business) is related to their pay, which lowers the analysts' incentive to reveal private information if it is not in favor of the employer's investment banking business.[45] However, it is very difficult to obtain data about analyst compensation; only case studies analyzing a single firm exist, and these are not able to provide a meaningful overview of the compensation plans of multiple brokers.

Overall, compared to the individual characteristics of analysts, the literature is largely silent on the influence of the institutional environment on analysts' herding behavior; in particular, there is no framework integrating both groups of factors. While broker size appears in the herding literature, the literature on analyst forecast errors mainly tackles the institutional factor of investment banking business. We thus consider these two factors in our analysis. Due to lack of data on analyst compensation, we are not able to include the institutional factor in our study, which is, unfortunately, usual in the literature. However, we aim to shed further light on the influence of the institutional environment on analyst herding behavior. Therefore, we address the issue of the lack of a general framework in the literature on analysts' herding behavior by explicitly combining institutional factors of the brokerage houses and the analysts' individual characteristics.

[44] See also, e.g., Malmendier and Shanthikumar (2007) and Hilary and Hsu (2013) for a comparison of the incentives and intentions of analysts to distort their forecast and not simply to provide the most accurate forecast possible.

[45] Following this reasoning, the Sarbanes-Oxley act and the SEC's Global Research Analyst Settlement aim at prohibiting the linkage of analysts' pay with investment banking activities.

4.2.3. The Costs of Non-Herding Framework

To capture the influence of both individual and institutional characteristics on analysts' forecasting behavior, we introduce a simple framework. We want to know which factors support analysts in conveying their private information and thus exhibiting non-herding behavior. Thus, we subsume all incentives that hinder the analysts' independence to reveal their private information and that make it costly for the analysts to engage in non-herding behavior (i.e. bold behavior) into the costs of non-herding (CONH). The CONH stand for all different kinds of costs, such as monetary, psychological or reputational costs stemming from the effort to establish a bold opinion in the brokerage (or the market) and from the resulting deterioration of career prospects.[46] Based on these terms, we can summarize the discussed individual and institutional factors affecting the CONH.

The literature states two individual factors that influence the CONH[47]: first, the analysts' individual reputation, and, second, the analysts' self-assessed ability. For the analysts' reputation, prior research shows empirically that a high reputation lowers, ceteris paribus, the analyst's CONH and, thus, facilitates bold behavior (see, e.g., Hong et al. 2000; Phillips and Zuckerman 2001; Clement and Tse 2005; Dyck et al. 2010). We also subsume the analyst's experience under the individual factor reputation. Experience is mostly another proxy for reputation, because experience documents analysts' track record and their survival in the highly competitive market, both indicating ability proxied by reputation. Clement and Tse (2005) show that both an analyst's experience with a certain stock (firm) and overall experience on the job are positively correlated with the analyst's probability of issuing a bold forecast, indicating a negative association with the CONH, in line with reputation.[48]

Second, self-assessed ability is negatively correlated with analysts' CONH (Clement and Tse 2005). A lower self-assessed ability increases the expected CONH as it raises the probability of a large forecast error and, thus, makes the analyst more prone to herd behavior.

[46] For example, in Bloomberg Markets Magazine, three equity analysts tell their stories of being bullied or even dismissed by their internal and external peers, sales peoples and supervisors because they forecast a bold (negative) economic trend in their industry sectors; finally, they quit their jobs. Ex post, all of them executed a correct analysis (Robinson 2009).

[47] We abstract from individual differences between analysts at the psychological level. Some analysts might have different dispositions and preferences with regard to risk behavior or other types of behavioral differences. For a comprehensive survey of behavioral differences, see Daniel et al. (2002).

[48] One could argue that, in general, older people, who have worked for a long time in a well-paid job such as that of a financial analyst, face lower CONH due to their wealth and resulting financial independence. However, we do not take such characteristics into account in this study.

Prior literature also points at other factors that affect analysts' herding behavior and that are usually seen as individual characteristics (see, e.g., Clement and Tse 2005 or Hilary and Hsu 2013). Various studies show that the *frequency* with which the analyst issues forecasts, the *days elapsed* since the last forecast has been issued, and the *number of days* remaining until the end of the stock's financial period are all positively correlated with the analyst's probability of issuing a bold forecast (see, e.g., Clement and Tse 2005). However, we do not identify these factors as either individual or institutional factors, nor as factors affecting the CONH of analysts directly. Instead, we classify them as factors that result from the individual and institutional determinants of analysts' forecasting behavior. For example, specific analysts have a low frequency of issuing forecasts because they already enjoy high reputations and, hence, do not need to show their ability by processing all incoming information and updating their forecasts (see, e.g., Hong et al. 2000).

On the other hand, institutional factors are also expected to influence the costs of non-herding. We are aware of four institutional factors cited in the literature as influencing analysts' forecasting behavior. The first is the way the analysts' performance is measured and compensated by the brokerage. As analysts' performance is usually evaluated by using a relative measure (benchmarking), the analysts face higher expected CONH, the more they deviate from the benchmark (Scharfstein and Stein 1990). This is because analysts who are close to the benchmark will not be blamed for anything with certainty; they can hide in the herd. The benchmarking incentivizes only highly able analysts to deviate from the herd (i.e. the consensus), because only they have a positive expected utility when deviating (see, e.g., Zwiebel 1995). In addition, as prior research has shown, long-term career concerns provide incentives for analysts to issue accurate forecasts (Mikhail et al. 1999 or Hong and Kubik 2003). Short term, during their appointment with a certain broker, analysts' compensation is usually not linked with the accuracy of their forecasts but with their ability to support their employer's investment banking business (Groysberg et al. 2011). Again, this factor raises the short-term CONH of analysts as they cannot independently decide to reveal their private information, because it will be scrutinized by investment banking executives. This shows the analysts' dilemma between having short-term costs with certainty when issuing bold forecasts or gaining long term benefits if their bold forecasts prove to be correct ex post and thus boost their career prospects.

Second, if analysts' abilities to support the investment banking business of their employer are positively linked to their compensation and (in-house) reputation, it seems obvious that the broker's investment banking business influences their forecasting behavior. Thus, we expect that the higher the broker's involvement in the investment banking business, the higher is the analyst's CONH.[49] Importantly, prior research is unclear about the effect of investment banking business (see, e.g., Cowen et al. 2006 or Barber et al. 2007). A larger involvement in investment banking business could also lead to a better access to information about the companies covered or could provide higher reputational costs for a brokerage house for issuing forecasts of low accuracy. Both effects might adjust or even exceed the negative effect of the broker's investment banking business on the analyst's CONH.

Third, the reputation of the brokerage house, instead of the reputation of the individual analyst, can support independent behavior and lower the CONH for the analyst, too. Both institutional and individual reputations are highly correlated because analysts with a high reputation are actively sought by first-tier brokerage houses (Hong and Kubik 2003).

Fourth, as Clement (1999) shows, a lower level of available resources increases the chance of issuing a less accurate forecast. When working in a small brokerage, the scarcity of resources raises the probability of missing or not receiving important information or of processing the information too slowly. Thus, analysts working in a scarce-resource environment face higher CONH, because they have to exert a higher effort to establish a bold opinion in the market and not to go with the herd.

The framework of the CONH provides a systematic way to disentangle and compare the effects of the various individual and institutional factors influencing analysts' forecasting behavior. In a next step, we introduce an additional institutional factor in order to make the framework on analysts' forecasting behavior more comprehensive.

4.2.4. The Ownership Effect and Analysts' Herding Behavior

Based on the framework of the CONH, we introduce an additional institutional factor that is expected to have a substantial impact on analysts' forecasting behavior but has so far not been considered in the literature. The new factor addresses a different channel than earlier determinants of analysts' forecasting behavior and focuses on one of the main differences of

[49] For example, Barber et al (2007, p. 493) state: "Perhaps the best-known example is Sanford Bernstein, which has been described as 'one of the more independent research houses – it only has a small syndicate business' (Financial Times, September 8, 2003, p. 26)." Accordingly, they provide evidence that unaffiliated analysts issue more accurate forecasts.

firms, namely their financing. As Rajan (2012) states in his *AFA* Presidential Address "The nature of firms and financing are intimately linked".

We will exploit the channel of the brokers' financing, thereby exploring the distinctive differences between privately held and publicly listed brokerage houses and their impact on analysts' CONH. Privately held brokerage houses are typically lower tier, employ a smaller number of employees, and are held by majority shareholders, who often participate in the management (Beatty and Harris 1999; Ke et al. 1999; Beatty et al. 2002). We expect the effect of the brokers' ownership on herding behavior to be twofold: (1) analysts from public brokers face higher CONH than analysts from private brokers when processing soft (i.e. bold) information and (2) this CONH is substantially larger for the subgroup of analysts from public brokers covering the financial sector, as their forecasts exert an externality on their employer's valuation.

(1): We rely on Stein's (2002) model on the information production in firms to explain theoretically why privately held brokerage houses might provide an institutional environment in which analysts face lower CONH and, hence, exhibit less herding behavior than analysts working in publicly listed houses. He explains theoretically how banks with flatter hierarchies and a more decentralized organizational setup are typically more efficient in producing and processing soft information, holding scope and size fixed. Soft information is costly to "harden" and cannot be easily transmitted inside the firm. Stein describes *soft information* as

> "[…] information that cannot be directly verified by anyone other than the agent who produces it. For example, a loan officer who has worked with a small-company president may come to believe that the president is honest and hardworking – in other words, the classic candidate for an unsecured "character loan." Unfortunately, these attributes cannot be unambiguously documented in a report that the loan officer can pass on to his superiors. This situation contrasts sharply with, for example, an application for a home mortgage loan. Here the decision of whether or not to extend credit is likely to be made primarily based on 'hard', verifiable information, such as the income shown on the borrower's last several tax returns." (Stein 2002, p. 1892)

A security analyst produces soft information when interpreting the existing information about a certain stock in an individual way, creating new information and, hence, a bold forecast. By doing so, the analyst cannot rely on the same interpretation of the facts as the majority of the analysts. Thus, bold analysts have to explain carefully and extensively (e.g., to their supervisors, investors or managers of the covered firm) why they are deviating from the opinion currently prevailing in the market. A stronger effort to communicate soft information results in higher CONH. Michaely and Womack (2005, p. 414) explain in detail how forecasts are "scrutinized by a research oversight committee and the legal department of the brokerage house before release." Based on such anecdotal evidence (see also Robinson 2009) and own interviews, we assume that publicly listed brokers have steeper hierarchies and more complex decision-making processes.[50] Therefore, analysts working in such a brokerage face higher costs in explaining, transmitting, and achieving bold forecasts.

Obviously, one could argue that banks without strong supervision of analysts' forecasting allow them to issue even more herding forecasts. Two considerations contradict this. First, such oversight committees have a hard job to establish their own opinion about the forecast the analyst wants to issue. Usually, the forecast is compared to other analysts' forecasts, which tends to make the analyst even more prone to herding behavior. Anecdotal evidence indicates that the broker's institutional incentives usually try to bias the analysts' forecasts toward the broker's interests as, for example, cited in Michaely and Womack (2005, p. 401) in the case of Morgan Stanley's internal memo: "Our objective […] is to adopt a policy, fully understood by the entire firm, including the Research Department, that we do not make negative or controversial comments about our clients as a matter of sound business practices". Second, as shown in prior studies, analysts issuing highly accurate forecasts face favorable career prospects (Hong and Kubik 2003). If issuing very inaccurate forecasts, analysts face high personal costs (turnover) and thus, are generally not motivated by individual career concerns to issue inaccurate forecasts *per se* (Hong et al. 2000; Groysberg et al. 2011).

There is also empirical evidence that supports Stein's (2002) model of soft and hard information in different organizational settings. Berger et al. (2005) confirm Stein's model by providing evidence that decentralized banks with short decision-making procedures are more able to collect, process, and act on soft information than centralized banks with tall hierarchies. Likewise, Rajan and Wulf (2006) and Guadalupe et al. (2012) show that there is a tendency to flatten the hierarchies in public firms, supporting Stein's notion (2002) that flat hierarchies deal better with soft information. Guadalupe et al. (2012) explicitly state the trend

[50] Please refer also to the interview study in Chapter 3. The interviews conducted with analysts that worked in privately held and publicly listed brokerage houses and top executives such as a former CEO of a Fortune 500 Bank in Switzerland confirm that privately held banks create a different environment to collect, transmit and process soft information, also in relation to analyst forecasts.

to organize publicly listed firms closer to partnerships, as the firms are facing a highly competitive environment with less physical capital intensive work; privately held brokerage houses usually constitute as partnerships (e.g., Rajan and Zingales 2001).

(2): The effect of ownership on analyst herding behavior is particularly pronounced for the subgroup of analysts covering their own sector, that is, the financial sector. By providing information about the condition of the firms they cover, analysts influence not only the valuation and the access to capital of the firms involved, but also (a) the firm's general reputation, its corporate strategy, and its compensation policies as well as (b) the reputation, compensation, and career prospects of the executives (Healy and Palepu 2001; Hribar and Jenkins 2004; Westphal and Clement 2008).[51] Therefore, the subgroup of analysts covering the financial sector while working at publicly listed brokerage houses do not only cover their employer's competitors and influence their valuation, but also the valuation of their own employer. By doing so, they also affect the reputation and compensation of their broker's executives. The ownership effect unveils the externalities of analysts' forecasts on their own employer's stock and all the related consequences, resulting in higher CONH for the analyst. This impact can be further rationalized by the fact that investors not only invest directly in specific stocks but also target stocks of a certain sector altogether and herd into certain industry sectors (see, e.g., Chevalier and Ellison 1999; Choi and Sias 2009). If, for example, analysts issue pessimistic earnings estimates about some of their employer's competitors, they can harm the valuation of their own employer and thus exert a negative externality on their listed brokerage house.

On the one hand, financial sector analysts from publicly listed brokerage houses have a lower incentive to reveal bad news and delicate information about their employers' competitors because they could negatively influence their own employers' valuation. On the other hand, financial sector analysts from a publicly listed brokerage do not have an incentive to disseminate bold positive news about a competitor either, as this could lead to a reduced valuation of their employer's stock compared to its competitor. However, financial sector analysts have to follow the information flow on the companies covered and ought to issue forecasts. Hence, the easiest way to issue a forecast without facing the costs of exposing

[51] This also explains why executives are interested in maintaining a positive relationship with analysts covering their firms. The literature shows that executives of stocks under coverage aim to influence the analysts to produce relatively positive forecasts about the future condition of the executives' firms. The analysts in turn aspire to receive favorable access to the top management of the firms they cover and other professional favors (Francis and Philbrick 1993; Lim 2001; Westphal and Clement 2008).

themselves to their executives or their employers' competitors is to hide in the herd. This reasoning explains why we are not interested in measuring a forecast error or a unidirectional bias of the analysts' earnings estimates but rather the herding behavior of analysts.

In contrast to publicly listed firms, analysts working with a privately held brokerage house encounter a fundamentally different and more independent institutional environment. The valuation of the analysts' employer does not depend on the stock market. Thus, they can reveal their private information to a larger extent independently than analysts working for publicly listed brokers.

This difference between financial sector analysts working at privately held and publicly listed brokers allows us to test the ownership effect across two parameters: Firstly via the higher costs of soft information production, in general; secondly, via the capital market dependency, in particular. While the first parameter provides the basis of the ownership effect equal among all analysts, the second one pronounces the effect for financial sector analysts. We expect that due to the capital market dependency, the difference in CONH between an analyst from a public broker and an analyst from a private broker is substantially larger if they cover stocks from the financial sector.

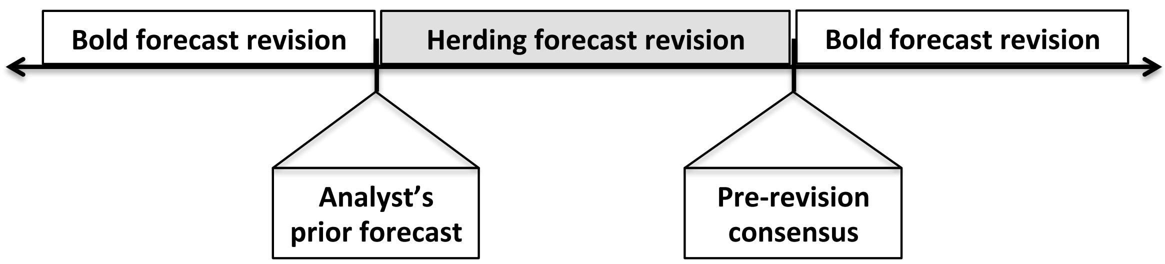
4.3. Data Sources, Variable Construction and Sample Selection

4.3.1. Institutional Factors and Boldness Measure

To test our hypothesis that the institutional factor of a brokerage house's ownership influences the forecasting behavior of an analyst, we employ data on analysts' earnings forecasts from the *Institutional Brokerage Estimation System* (*I/B/E/S*) database. We matched the analysts' earnings forecast data with the data on the ownership (i.e., privately held or publicly listed) of the brokerage house the analyst is working for. We used the *I/B/E/S* translation files to obtain the names of the brokerage houses, and we hand-collected data on their ownership for the entire sample period by using various sources: *FINRA-BrokerCheck* of the Financial Industry Regulation Authority, the *Institution Search* database of the National Information Center of the Federal Financials Institutions Examination Council (FFIEC), the *Investment Adviser Information Reports* of the Security Exchange Commission (SEC), *Bloomberg Businessweek*, *LexisNexis*, *Wikipedia*, and company homepages. Also, we cross-checked our data with the data available from the appendixes of prior studies that track changes of broker names over several years due to M&A activities (Corwin and Schultz 2005; Hong and Kacperczyk 2010; Bao and Edmans 2011). We checked possible affiliations in each year of the data sample and classified each brokerage house for each year into the binominal variable *Private,* taking the value of 1 if the analyst was working for a privately held brokerage house; and 0 otherwise. If a brokerage house was a subsidiary of a publicly listed brokerage house or a listed financial services company, we classified this broker as not privately held.

As our dependent variable, we construct a measure for herding behavior with the earnings estimates' data from the detailed history file of *I/B/E/S*. We focus on forecast revisions because before issuing a revision, the analyst *i* has to decide whether to follow the herd or to trust in his or her own private information (i.e., issuing a bold forecast). We use the binominal boldness measure (*Bold*) introduced by Gleason and Lee (2003) and Clement and Tse (2005) to distinguish between bold and herding forecast revisions (see Figure 1).

Figure 1: Measure of Analysts' Herding Behavior



An analyst deviating from both, his or her own prior earnings estimate *and* the prerevision consensus reveals new private information to the market and, thus, issues a bold forecast (*Bold* = 1). In contrast, forecast revisions are classified as herding forecasts (*Bold* = 0) if they lie between the analyst's prior forecast and the pre-revision consensus, as illustrated in Figure 1.[52] This captures the idea that the analysts exhibit herding behavior when deviating from their prior forecasts and revising towards the consensus. All other revisions that deviate from both the prior forecast and from the prerevision consensus are classified as bold, and the variable *Bold* takes on the value of 1. Prior research has documented that investors react more

[52] An additional detailed graphic explanation of the dependent variable measuring the herding behavior of an analyst is presented in Appendix B.

strongly to bold revisions than to herding revisions, providing evidence that bold revisions reveal more novel information to the market (Gleason and Lee 2003; Clement and Tse 2003).

To additionally control for the investment banking business of the analyst's brokerage house, we obtained data for the underwriting business from the database *SDC Platinum*. Again, as with the data for *Private*, we checked whether the broker was affiliated with a listed parent company and classified them according to their parent company. Similarly to Cowen et al. (2006) and Barber et al. (2007), we manually classified the various brokerage houses into four categories in each sample year: (1) *Lead_Underwriter*, brokerage houses (or their parent company) that served as lead underwriter in year *t* on at least one U.S. IPO or SEO[53]; (2) *Co_Underwriter,* brokerage houses (or their parent company) that acted as a co-manager in at least one U.S. IPO or SEO in year *t*; (3) *Syndicate_Member,* brokerages (or their parent company) that provided services as a syndicate member in at least one U.S. IPO or SEO in year *t*; (4) *No Underwriting,* brokerage houses (or their parent company) that were not registered in the SDC database and therefore are assumed not to have provided any underwriting service in year *t.*

To generate a measure of the analyst's resource environment and/or broker reputation, we relied on previous research and counted the number of analysts working with a specific brokerage house for each year (Clement 1999). We call this variable *BrokerSize* and use the log of (1+number of analysts) in our empirical analysis (see, e.g., Malloy 2005; Hilary and Hsu 2013).

4.3.2. Control Variables

To control for the analysts' individual effects on their forecasting behavior, we constructed a comprehensive set of analyst-firm variables per year, that is, for analyst *i* who issues a forecast for firm *j* in the year *t*. First, we controlled for the analyst's lagged accuracy (*LagAccuracy*), defined as the analyst *i'*s accuracy of last year (*t-1*) for stock *j*. We calculated the analysts' accuracy (*Accuracy*), defined as the maximum absolute forecast error (*AFE*)[54] for analysts covering firm *j* in year t ($AFEmax_{jt}$) minus the *AFE* of analyst *i* following firm *j*

[53] SEO stands for a seasoned equity offering and IPO for an initial public offering. An IPO indicates the going public of a formerly privately held company and the SEO a new equity issue of an already publicly listed company.

[54] The absolute forecast error (AFE) is defined as the absolute difference between the analyst's EPS estimate (F) and the firm's actual EPS (A) in year t scaled by the stock price (P) of firm j two days before the revision: AFEjit = |Fijt - Ajt| / Pj (see, e.g., Hong and Kubik 2003 or Clement and Tse 2005).

in year t (AFE_{ijt}), with this difference scaled by the range of the maximum ($AFEmax_{jt}$) to the minimum ($AFEmin_{jt}$) of all the *AFE* for analysts covering stock *j* in year *t* (see, e.g., Hong and Kubik 2003 and Clement and Tse 2005).[55] This measure takes on values between 0 and 1, where 1 is the most accurate estimate of an analyst covering stock *j* in year *t* and 0 is the least accurate estimate of an analyst covering the same stock *j* in the same year t (see equation 1):

$$Accuracy_{ijt} = \frac{AFE\max_{jt} - AFE_{ijt}}{AFE\max_{jt} - AFE\min_{jt}} \qquad (1)$$

Next, we followed previous research to control for other individual analyst-firm characteristics (e.g., Lim 2001; Hong et al. 2000; Hong and Kubik 2003; Jackson 2005; Clement and Tse 2005; Hilary and Hsu 2013). *DaysElapsed* counts the days elapsed since the last forecast by any other analyst following the same firm *j* in year *t*. *Horizon* measures the number of days between the forecast date and the end of the fiscal period. *Frequency* is the number of the forecasts the analyst *i* issued on firm *j* during year *t*. *GenExperience* is the general experience of analyst *i*, measured as the number of years analyst *i* was employed as analyst. *FirmExperience* is calculated as the number of years analyst *i* covered firm *j* in year *t*. Both variables measure the experience of the analyst that is calculated on the base of the maximal range of the *I/B/E/S* database, starting in 1983. As we did not want to lose any observations due to left-censored analysts nor introduce a possible selection bias by excluding them, we used the full *I/B/E/S* sample. *Companies* is the number of companies analyst *i* is covering in year *t*. Similarly, *Industries* counts the number of two-digit SIC codes the analyst *i* covers in year *t*.

[55] For example, imagine an analyst *i* who issues an EPS forecast of USD 2.50 for stock *j* in year *t*. The actual EPS of stock *j* is USD 3. Thus, analysts *i*'s absolute forecast error equals AFE_{ijt}=(|3-2.50|/100)= 0.005 as the stock price of *j* two days before analyst *i* issued the EPS forecast was USD 100. The most inaccurate analyst covering the same stock *j* bears a $AFEmax_{jt}$ of 0.02 because his or her EPS forecast was USD 1.00 and the most accurate analyst covering stock *j* yields $AFEmin_{jt}$ of 0.001 as the forecast was USD 2.90. Finally, the *Accuracy* of analyst *i* covering stock *j* is in year *t* is $Accuracy_{ijt}$= (0.02-0.005)/(0.02-0.001)=(0.015/0.019)=0.789. If analyst *i* were the most accurate analyst in this example, his *AFE* would be equal to *AFEmin*, resulting in an *Accuracy* of 1, which is naturally the highest value for *Accuracy*.

4.3.3. Sample Selection

We obtained the data on analysts' earnings estimates (EPS) from the *I/B/E/S Detailed Earnings History Files.* For the translation of the *I/B/E/S* broker codes into the names of the brokerage houses, we used the *I/B/E/S Broker Translation Files*. As *I/B/E/S* (Thomson Reuters) ceased providing the Broker Translation File in late 2009, our sample period ends in 2008.[56] Our data starts in 1999 and spans a range of 10 years so as to have enough observations to provide statistically meaningful hypothesis tests, similar to previous research (see, e.g., Malloy 2005; Clement and Tse 2005; Malmendier and Shanthikumar 2007).

In accordance with prior studies focusing on annual earnings estimates, we retained the last earnings estimate an analyst i issued for a specific stock j in a particular year t, at least 30 days and at most 1 year prior to the end of stock j's financial period (O'Brien 1990; Sinha et al. 1997; Clement 1999; Clement and Tse 2005). The idea behind the focus on an analyst's last EPS revision of a stock in a financial year is to achieve a homogenous group of forecasts, as analysts can revise their forecasts arbitrarily and in variable intervals. This leads to a reduction by around 75% of observations of the original data set, as an analyst typically issues about 4 estimates per stock per year. As we compare forecasts of a specific stock in a specific year, we had to eliminate all observations for which only one forecast per stock-year exists and observations with no prior year data on the analyst's forecast accuracy. In addition, we excluded observations with no data on the ownership of the analyst's brokerage house.[57]

Further, we followed Clement and Tse (2005) and eliminated potential outliers in four steps: First, we obtained stock prices from *The Center For Research in Security Prices (CRSP).* Second, we facilitated comparisons across stocks by deflating forecast revisions and forecast errors (forecasted EPS – actual EPS) by the firm's stock price 2 days before the forecast revision date. Third, we omitted outliers with these price-deflated forecast revisions above 0.10 or below -0.10 and discarded observations with a price-deflated analyst error above 0.40 or below -0.40.[58] This resulted in an unbalanced panel data set of 136,428 earnings estimates (observations) issued by 5,760 analysts working with 228 brokerage houses during the years 2000 to 2008.

---

[56] Note that because our analysis focuses on the last EPS forecast the analyst issues at least 30 days before the end of the reporting period, the year 2008 is the last possible sample period.

[57] Due the restriction of missing ownership data, we lose 7.8% of the final observations.

[58] Clement and Tse (2005) argue that these outliers are potentially distressed firms, suggesting that these observations are not comparable to the usual forecasts of analysts. In the same vein, other studies exclude observations in which the stock price is less than five dollars, or forecasts whose absolute forecast error exceeds 10 dollars (see, e.g., Hong and Kacperczyk 2010).

4.4. Methods and Predictions

We regress our measure for herding behavior (*Bold*) on our institutional factor of ownership (*Private*), controlling for other firm- and analyst-specific differences, as described in our framework. As *Bold* is a binominal variable, we use a non-linear logit model. We begin with the herding behavior model by Clement and Tse (2005), which controls for a full set of individual analyst characteristics, and extend it by adding the explanatory variable *Private* and the interaction term between *Private* and *BrokerSize* (*INT_BrokerSize_Private*) as well as controls for the broker's underwriting business as specified in regression equation (2):

$$\text{Logit}(\text{Bold}_{ijt}) = \alpha_0 + \alpha_1 \text{Private}_{ijt} + \alpha_2 \text{BrokerSize}_{ijt} + \alpha_3 \text{INT\_BrokerSize\_Private}_{ijt} + \alpha_m \text{SDC}^m_{ijt} + \alpha_k X^k_{ijt} + \varepsilon_i \qquad (2)$$

where Bold is the binominal boldness measure for analyst i's earnings estimate for firm j in year t; Private is the binominal (privately held or publicly listed) ownership variable of the brokerage house the analyst i covering stock j is working with in year t; BrokerSize is the log of (1+number of analysts) working with analyst i covering stock j in year t; INT_BrokerSize_Private is the interaction term; SDC is the vector comprising m broker-specific control variables for the brokerages' underwriting activities the analyst i is working with (Lead_Underwriter, Co_Underwriter, Syndicate_Member, No_Underwriting);[59] and the vector X subsumes k individual control variables for analyst i covering firm j in year t (LagAccuracy, Horizon, DaysElapsed, Frequency, GenExperience, FirmExperience, Companies, and Industries). Based on equation (2) we state our predictions as follows:

Prediction 1: We expect the coefficient α_1 of the independent variable Private to have a negative impact on the analyst's probability of issuing a bold forecast (i.e. $0 > \alpha_1$).

We expect a negative basic impact on the analyst's probability of issuing a bold forecast if he or she is employed at a privately held broker because private brokers have on average fewer informational resources and less reputation to establish an autonomous bold view in the market, which increases their analysts' CONH. Although we assume that the analysts from

[59] In the regression, we include binary variables for *Lead_Underwriter, Co_Underwriter, and Syndicate_Member; No_Underwriting is the base category.*

private brokers process soft (i.e. bold) information in a less costly way (lower CONH), we predict a negative general effect of *Private* compared to analysts from public brokers, even for public brokers of the same size. This is because smaller public brokers are affiliated with larger parent companies which provide them with access to more relevant information than unaffiliated, privately held brokers of the same size.

Prediction 2: We expect the coefficient of BrokerSize α_2 to have a positive impact on the analyst's probability of issuing a bold forecast (i.e. $0 < \alpha_2$).

The positive basic impact on the analyst's probability of issuing a bold forecast is shown in previous research. Hence, more resources imply lower CONH for analysts, resulting in a higher probability of issuing bold forecasts.

Prediction 3: We expect the coefficient α_3 of INT_BrokerSize_Private to have a positive impact on the analyst's probability of issuing a bold forecast and that this effect is larger than the effect of BrokerSize (i.e. $0 < \alpha_2 < \alpha_3$).

For the interaction term *INT_BrokerSize_Private*, we state a similar prediction as for *BrokerSize* but expect a larger coefficient because, firstly, we assume that the analysts from private brokers process soft (i.e. bold) information in a less costly way than analysts from public brokers; secondly, the larger the private broker becomes, the smaller is the difference in informational resources. Analysts from smaller brokers that are affiliated with a public broker (and, thus, are naturally coded as *public* brokers) receive access to larger informational resources than analysts from unaffiliated privately held brokers of the same small size. Hence, the growth in size increases the probability of issuing a bold forecast *more* for analysts from a private broker than for analysts from public brokers.

Prediction 4: *We expect the overall effect of private ownership (Private and INT_BrokerSize_Private combined) on the analyst's probability of issuing a bold forecast to become positive after the privately held brokerage reaches a certain level of BrokerSize.*

We reason that the threshold lies around the number of 25 analysts employed at a broker, as previous research has shown this number to define a crucial size for a brokerage, both for its resources and for its reputation (see, e.g., the definition of high status brokerage in Hong et al 2000). However, compared to Hong et al. 2000, the average *BrokerSize* increased during our sample period (1999-2008). Thus, we expect a higher threshold by trend.

We have no clear prediction for the coefficient α_m of the control variables for the underwriting business of the broker included in the vector *SDC*; prior research remains unclear about the effect of the broker's investment banking business on the CONH of its analysts. We expect the coefficients α_k of the individual characteristics captured in the vector X to behave as shown in prior research studies (see, e.g., Clement and Tse 2005).

To test whether the effect of ownership on analyst herding is more pronounced for analysts covering their own sector (i.e. the financial sector), we split the sample into two subsamples, (A) analysts covering the financial sector, and (B) all other analysts.[60] The split is executed according to the two-digit SIC codes of the stocks each group covers as used by the U.S. Department of Labor (Division H: Finance, Insurance and Real Estate); it allocates all SIC codes from 6000 to 6999 into (A) Financial sector stocks, and all other codes into (B).[61]

We regress the model from equation (2) for both sub-samples on the basis of analyst-stock observations and compare the resulting coefficients. We expect the difference in CONH, and thus in herding behavior between analysts at privately held brokerage houses and analysts at publicly listed houses, to be substantially larger for analysts covering the financial sector than for analysts covering other sectors. This difference is captured for both sub-samples in the coefficients of the explanatory variables *Private* (α_1) and the interaction term *INT_BrokerSize_Private* (α_3).

4.5. Results

4.5.1. Descriptive Statistics

Panel A of Table 1 shows the descriptive statistics for the full sample of our study. Most analysts make bold forecast (78%), consistent with their incentives to signal their release of new private information to the market. This finding is in line with prior research about the herding behavior of stock analysts. Gleason and Lee (2003) find 76% of forecasts in their sample are "high-innovation" (i.e. bold), while Clement and Tse (2005) find about 74%. However, these figures do not signal that there is no problem with the herding behavior of

[60] Security analysts mostly cover only one sector as they need highly specialized industry expertise.

[61] The SIC codes industry classification of the Department of Labor is accessed through their homepage (http://www.osha.gov) on June 16, 2012. This classification is identical to the Fama-French industry portfolios (industry portfolio number 12: *Finance*) accessible through Kenneth French's page (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french); accessed on June 16, 2012.

financial analysts, as about a quarter of all forecasts indicate herding behavior, not revealing private information to the market. That is why we want to analyze whether the institutional factor of ownership systematically hinders analysts from revealing their full private information, i.e. to issue a bold forecast.

In the full sample, 30,437 (22%) firm-analyst observations stem from analysts who are employed at private brokerage houses. On average, an analyst works for a brokerage that employs about 67 analysts. The interquartile range lies between 24 and 109 analysts.[62] The most firm-analyst observations, 105,991 (77%), come from analysts working with brokerage houses that act as lead underwriters, 10% of the observations come from brokers acting as co-underwriters, 3% from analysts at syndicating brokers, and 9% from analysts working with brokerage houses that are, according to the SDC database, not involved in underwriting business. The average analyst has a general experience of 6 years (*GenExperience*), has experience per covered firm of almost 4 years (*FirmExperience*), covers about 17 different companies (*Companies*), and covers almost 4 different industries (i.e. two-digit SIC sectors, *Industries*). On average, the analysts issue their forecasts 13 days after another analyst covering the same stock issues a forecast (*DaysElapsed*), they issue almost 5 forecasts per covered firm per year (*Frequency*), and they do this 95 days before the firm's financial period ends (*Horizon*: remember that we restricted the observations to the first 11 months of the firm's financial period). Most of the analysts make accurate forecasts with some analysts issuing very poor forecasts, indicated by the substantially skewed distribution for lagged accuracy (*LagAccuracy*, mean 0.73, median 0.85).[63]

Panel B of Table 1 shows the subsample for forecasts of stocks of the financial sector. The subsample is very similar to the full sample. The main variables are identical in their distribution (e.g., *Bold*) or differ only marginally (e.g., *BrokerSize*). Compared to the full sample, the subsample yields only 3% more observations that come from analysts at privately held brokerage houses (total 25%). In addition, the individual analyst characteristics have an almost identical distribution as in the full sample and in the distribution of the brokerage houses' underwriting business variables. The only considerable difference stems from the

[62] Note that *BrokerSize* here shows the raw number of analysts employed at a brokerage and not the log of (1+number of analysts); this is to provide a meaningful statistic.

[63] The lagged accuracy is scaled between 0 and 1, where 0 is the forecast with the largest absolute forecast error of an analyst *i* covering stock *j* in year *t-1* and 1 defines the best forecast of another analyst *k* also covering stock *j* fin year *t-1*.

distribution of the number of industries the analyst covers, which is simply explained by the restriction to forecasts on stocks from the financial sector.

Panels C and D of Table 1 show the distribution of the different variables for the division of the full sample into observations from analysts working at privately held brokerage houses and from analysts working at publicly listed houses, respectively. There are no substantial differences for these two subsamples in the individual analyst's forecasting characteristics, as in *Horizon* or *Frequency*, whereas the privately held houses on average employ slightly fewer experienced analysts. Likewise, analysts at privately held houses cover slightly fewer *Companies* but more *Industries*.

However, the two subsamples differ substantially across their institutional environment. Privately held brokerage houses employ on average far fewer analysts (22) and are not as involved in underwriting activities as publicly held houses. Only 51% of the observations stem from privately held houses acting as lead underwriters, but 22% come from brokers acting as co-underwriters, 11% as syndicate member, and 16% from privately held houses that are not involved in underwriting activities at all. In contrast, publicly listed houses on average employ almost 80 analysts. Furthermore, a larger number of the analysts from public brokers are involved in underwriting business, as expected. Some 85% of all observations come from publicly listed houses acting as lead underwriters, and only 7% from houses operating as co-underwriters, almost no observations from syndicate member houses (1%), and again, only 7% of all observations stem from houses not involved in underwriting business.[64]

Panels E and F of Table 1 (see Appendix B, Table B2) manifest the same differences between the two subsamples, although restricted to forecasts of stocks from the financial sector. The differences in *BrokerSize* and the underwriting business of the diverse brokerage houses remain. Again, the forecasting behavior (*Bold*) does not differ in essence across private and public brokers. This univariate finding reflects the trade-off between the size of private brokers and their independence, affecting the analysts' CONH, and, hence, the likelihood of bold forecasts. Private brokers are, on average, smaller and thus have fewer informational resources available, increasing the CONH and reducing the probability of their analysts issuing bold forecasts. However, analysts from private brokers face lower CONH and, thus, might have a higher probability of issuing a bold forecast.

[64] This is usually the case for smaller U.S. broker that are subsidies of larger foreign banks that are not involved in the U.S. IPO or SEO underwriting market.

Table 1: Descriptive Statistics of Individual Analyst and Institutional Broker Characteristics

Panel A: Full Sample (N=136,428)						
Variables	Mean	Min	25th	Median	75th	Max
Bold	0.78	0	1	1	1	1
Private	0.22	0	0	0	0	1
BrokerSize	66.69	1	24	51	109	229
LagAccuracy	0.73	0	0.59	0.85	0.97	1.00
DaysElapsed	13.33	1	1	5	16	90
Horizon	95.41	31	59	69	110	357
Frequency	4.80	2	3	4	6	33
GenExperierience	6.07	0	3	5	8	26
FirmExperience	3.79	0	1	3	5	26
Companies	17.11	1	12	16	20	104
Industries	3.87	0	2	3	5	22
Lead_Underwriter	0.77	0	1	1	1	1
Co_Underwriter	0.10	0	0	0	0	1
Syndicate_Member	0.03	0	0	0	0	1
No_Underwriting	0.09	0	0	0	0	1

Panel B: Financial Sector Coverage (N=21,823)						
Variables	Mean	Min	25th	Median	75th	Max
Bold	0.77	0	1	1	1	1
Private	0.25	0	0	0	0	1
BrokerSize	58.87	1	21	38	85	229
LagAccuracy	0.71	0	0.56	0.84	0.97	1.00
DaysElapsed	13.22	1	1	5	16	90
Horizon	94.74	31	60	70	103	348
Frequency	4.55	2	3	4	6	21
GenExperience	5.83	0	3	5	7	26
FirmExperience	3.63	0	1	3	5	24
Companies	18.92	1	13	18	23	87
Industries	2.83	1	2	2	3	22
Lead_Underwriter	0.75	0	1	1	1	1
Co_Underwriter	0.13	0	0	0	0	1
Syndicate_Member	0.04	0	0	0	0	1
NoUnderwriting	0.08	0	0	0	0	1

Panel C: Privately Held Brokers (N=30,437)						
Variables	Mean	Min	25th	Median	75th	Max
Bold	0.78	0	1	1	1	1
BrokerSize	21.67	1	11	20	29	61
LagAccuracy	0.71	0.00	0.54	0.84	0.97	1.00
DaysElapsed	15.29	1	1	5	19	90
Horizon	93.87	31	59	68	108	346
Frequency	4.60	2	3	4	6	21
GenExperience	5.63	0	3	5	7	26
FirmExperience	3.47	0	1	2	4	26
Companies	16.77	1	12	15	20	78
Industries	4.37	0	2	4	6	22
Lead_Underwriter	0.51	0	0	1	1	1
Co_Underwriter	0.22	0	0	0	0	1
Syndicate_Member	0.11	0	0	0	0	1
No_Underwriting	0.16	0	0	0	0	1

Continuation of Table 1

Panel D: Publicly Listed Brokers (N=105,991)						
Variables	Mean	Min	25th	Median	75th	Max
Bold	0.78	0	1	1	1	1
BrokerSize	79.62	1	35	63	121	229
LagAccuracy	0.73	0.00	0.60	0.85	0.97	1.00
DaysElapsed	12.76	1	1	4	15	90
Horizon	95.85	31	59	69	111	357
Frequency	4.86	2	3	4	6	33
GenExperience	6.19	0	3	5	8	26
FirmExperience	3.88	0	1	3	5	26
Companies	17.21	1	12	16	21	104
Industries	3.72	0	2	3	5	22
Lead_Underwriter	0.85	0	1	1	1	1
Co_Underwriter	0.07	0	0	0	0	1
Syndicate_Member	0.01	0	0	0	0	1
No_Underwriting	0.07	0	0	0	0	1

Notes: The table reports summary statistics for the full sample (Panel A), the subsample of analysts covering the financial sector (Panel B), the subsamples of analysts from privately held (Panel C) and publicly listed brokers (Panel D), and the subsamples of analysts covering the financial sector from privately held (Panel E) and publicly listed brokers (Panel F). We retain only the last forecast an analyst *i* is issuing on a specific stock *j* in year *t*. The variable *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). *Private* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the number of analysts employed at a specific broker. *LagAccuracy* is analyst *i*'s last year accuracy on stock *j*. *DaysElapsed* is the number of days that are elapsed between analyst *i*'s forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst *i*'s forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst *i* makes during the financial year *t* for a stock *j*. *FirmExperience* is the number of consecutive years analyst *i* covers stock *j*. *GenExperience* is the number of consecutive years in which analyst *i* filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst *i*. *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year *t* as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO. *No_Underwriting* equals 1 if the brokerage house is not engaged in underwriting activities in year *t*.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia and company home pages.

4.5.2. Regression Results

Table 2 shows five models on the basis of regression equation (2). The first, model (1), includes only the individual analyst characteristics and one institutional factor, *BrokerSize,* identically to Clement and Tse (2005); the last, model (5), incorporates the ownership variable *Private*, the interaction term between *Private* and *BrokerSize*, and the full set of controls as specified in regression equation (2). Models (2) to (5) show the development of the coefficients when adding our explanatory variables of interest (*Private* and *INT_Private_BrokerSize*) and other controls.

Model (2) includes the ownership variable *Private*. The coefficient for *Private* is positive (0.031) and statistically significant at the 10% level, suggesting that analysts working with privately held brokers are, on average, more likely to issue bold forecasts. The coefficient of *BrokerSize* is positive and highly significant, indicating that analysts from larger brokers with more available resources have a higher probability of releasing bold forecasts.

Model (3) includes the interaction effect between *Private* and *BrokerSize*. The positive (0.066) and significant effect of *INT_Private_BrokerSize* and the negative (−0.174) and significant effect of *Private* show that privately held brokers face a trade-off: On the one hand, privately held brokers on average have fewer resources available, leading to higher CONH and fewer bold forecasts; on the other hand, analysts working with privately held brokerage houses face lower costs of processing soft information, resulting in lower CONH. If the number of analysts employed at a private broker increases, the resources available increase. Similar to model (2), *BrokerSize* is still positive and significant in model (3); thus, an additional analyst lowers the analysts' CONH and increases the probability of issuing a bold forecast for publicly listed brokers, too. However, in model (3), the coefficient is almost three times smaller for public brokers than for private brokers.

We analyze the average marginal effect (AME) of a change in ownership *(Private)* in model (3) by calculating the predictive probability for all observations and averaging the effect of working for a privately held or for a publicly listed broker. This procedure yields a positive AME of 1.7% for *Private*, highly significant at the 1% level, indicating a higher probability of issuing a bold forecast for an analyst working with a private broker than for the same analyst working with a public broker.[65] In model (4), we include the set of underwriting variables as additional institutional controls. First and foremost, the main variables of interest, *Private* and *INT_Private_BrokerSize*, remain constant in both magnitude and significance, confirming that the ownership effect is distinctly different from the effect of the broker's underwriting business. In contrast, the coefficient of *BrokerSize* drops substantially and loses statistical significance.

[65] The standard errors are calculated with the delta method to account for the finite, discrete change of *Private* (0 to 1). We also conduct an analysis of the discrete probability effect (DPE) of being with a privately held or publicly listed brokerage house calculated at the means of the variables in the model. The analysis yields that the average analyst working with a privately held house has a 1.8% higher probability of issuing a bold forecast than the same analyst with a public broker. The effect is significant at the 1% level. The similar result of the DPE to the AME supports the robustness of the effect.

Table 2: Regression of Forecast Boldness on Institutional Factors and Analyst Characteristics (Full Sample)

Dependent variable: *Bold*					
Variables	(1)	(2)	(3)	(4)	(5)
Private		0.0311*	-0.1743**	-0.1712**	-0.1712**
		(1.68)	(-2.36)	(-2.32)	(-1.99)
INT_Private_BrokerSize			0.0663***	0.0662***	0.0662**
			(2.87)	(2.87)	(2.15)
BrokerSize	0.0262***	0.0333***	0.0235***	0.0067	0.0067
	(3.68)	(4.03)	(2.62)	(0.67)	(0.41)
LagAccuracy	0.1148***	0.1153***	0.1153***	0.1144***	0.1144***
	(5.39)	(5.41)	(5.41)	(5.36)	(5.45)
DaysElapsed	0.0041***	0.0040***	0.0040***	0.0040***	0.0040***
	(11.49)	(11.43)	(11.37)	(11.19)	(11.51)
Horizon	0.0004***	0.0004***	0.0004***	0.0005***	0.0005***
	(3.56)	(3.61)	(3.61)	(3.81)	(3.14)
Frequency	-0.0005	-0.0005	-0.0004	-0.0002	-0.0002
	(-0.18)	(-0.15)	(-0.14)	(-0.06)	(-0.04)
GenExperience	0.0039	0.0040*	0.0041*	0.0036	0.0036
	(1.62)	(1.65)	(1.70)	(1.50)	(1.16)
FirmExperience	-0.0025	-0.0024	-0.0022	-0.0016	-0.0016
	(-0.81)	(-0.79)	(-0.70)	(-0.51)	(-0.46)
Companies	-0.0058***	-0.0058***	-0.0058***	-0.0059***	-0.0059***
	(-6.99)	(-6.90)	(-6.96)	(-7.04)	(-4.12)
Industries	0.0159***	0.0155***	0.0155***	0.0165***	0.0165***
	(5.41)	(5.28)	(5.27)	(5.58)	(3.63)
Lead_Underwriter				0.1448***	0.1448**
				(6.13)	(2.06)
Co_Underwriter				0.1123***	0.1123*
				(3.82)	(1.68)
Syndicate_Member				0.0916**	0.0916
				(2.21)	(1.15)
Constant	1.0107***	0.9750***	1.0146***	0.9491***	0.9491***
	(23.82)	(20.53)	(20.51)	(18.01)	(11.19)
Clustered by brokerage house	No	No	No	No	Yes
Observations	136,428	136,428	136,428	136,428	136,428
Chi2	245.0	248.0	255.1	292.0	306.9

Notes: The table reports logit coefficient estimates, and in parentheses, robust z-statistics. In model (5), standard errors are clustered by broker. We retain only the last forecast an analyst i is issuing on a specific stock j in year t. The dependent variable *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). *Private* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the natural logarithm of the number of analysts employed at a specific broker. *INT_Private_BrokerSize* is the interaction term between *Private* and *BrokerSize*. We control for the analyst's individual characteristics: *LagAccuracy* is analyst i's last year accuracy on stock j. *DaysElapsed* is the number of days that are elapsed between analyst i's forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst i's forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst i makes during the financial year t for a stock j. *FirmExperience* is the number of consecutive years analyst i covers stock j. *GenExperience* is the number of consecutive years in which analyst i filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst i. In models (4) and (5), we also include controls for underwriting activities: *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year t as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia and company home pages.

Because observations from analysts of the same brokerage house are nested in model (4), we cluster the standard errors at the brokerage level in model (5).[66] The results from model (5) with clustered standard errors confirm the ownership effect. However, due to the nonlinearity of our model, we should interpret marginal effects with caution, particularly as we use interaction effects.[67] Again, we first conduct an analysis of the AME of a change in ownership *(Private)* in model (5), yielding a positive difference of 1.8% for analysts working with private brokers compared to the same analysts working with public brokers, significant at the 10% level.[68] So far, analysts from private brokers face lower CONH on average and, hence, issue more bold forecasts revealing new private information to the market.

To provide a more insightful analysis of the ownership effect, we calculate the marginal effects of *Private* restricted to the common support for *BrokerSize*, which ranges from 0 to 61 analysts, as the largest privately held broker employs 61 analysts. The marginal effects of ownership for model (5) are plotted in Figure 2. We show the AME conditional on *BrokerSize* and determine the underwriter level to be the highest (i.e., *Lead_underwriter* = 1), because the majority of the observations of both ownership types belong to this group.[69]

Panel A in Figure 2 presents a steeper slope of *BrokerSize* for privately held brokerage houses, indicating a larger positive marginal effect of *BrokerSize* when the broker is privately owned. In contrast, for publicly listed brokers, there is virtually no marginal effect of an additional analyst employed, indicating diminishing marginal returns and the existence of a maximum for the impact of informational resources on the probability of issuing bold forecasts.

---

[66] We cluster at brokerage level, as we are interested in the effect of institutional factors idiosyncratic to all the observations of all the analysts working with the same specific broker. As the brokerage level is superior to a clustering at the analyst level, we apply the more stringent test for significance. When we cluster at the analyst level, we arrive at smaller standard errors for all the main coefficients.

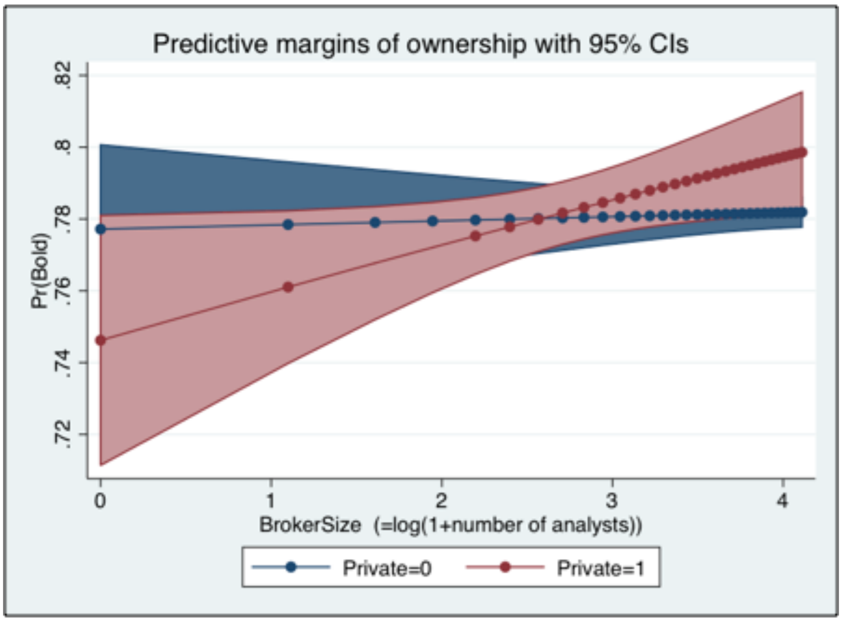
[67] We acknowledge possible problems concerning the interpretation of marginal effects in non-linear models as well as the additional complicacy with regard to interaction effects. Please refer to our robustness section, where we conduct additional tests to address such concerns more extensively.

[68] The difference in significance of the AME between model (4) and model (5) stems mainly from the clustering of the standard error at the brokerage level. The analysis of the DPE at means of being with a privately held or publicly listed brokerage house yields that the average analyst working with a privately held house has a positive difference of 1.4 percentage points in the probability of issuing a bold forecast compared to the same analyst from a public broker. The effect is significant at the 10% level. The similar result to the AME again supports the robustness of the effect.

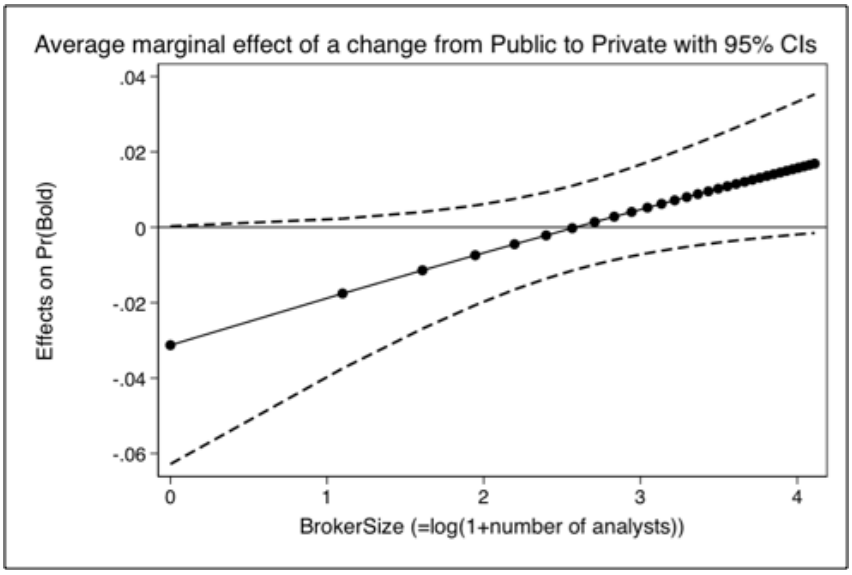
[69] All other underwriting variables (Co_Underwriter, Syndicate_Members and No_Underwriting) are kept at zero. This procedure is also called marginal effects at representative values (MERs).

Figure 2: Plotted Marginal Effects of *Private* for Different *BrokerSizes* (Full Sample)

Panel A

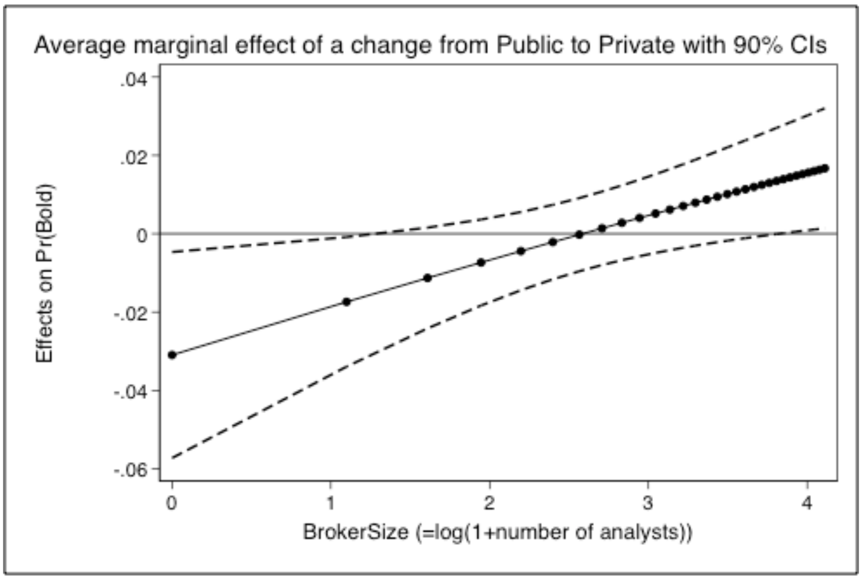


Panel B



(Continuation of Figure 2)

Panel C



The difference between the marginal effects of *BrokerSize* for the two ownership types of brokerage houses, as shown in Panels B and C of Figure 2, has its break-even point around 12 analysts (which corresponds to a *BrokerSize* of 2.5 on the x-axis of the figures).[70] At the 5% significance level, the difference does not attain significance over the range from 0 to 61 analysts, as shown in Panel B. At the 10% level in Panel C, the difference becomes significant around 40 analysts.[71] Thus, analysts working for privately held brokers that employ more than 40 analysts show a significantly higher probability of issuing a bold forecast than a similar analyst at a listed broker. At this threshold, the difference in probability of around 2.6% predicts an economically meaningful effect of ownership on analysts' forecasting behavior in relation to the large number of forecast revisions issued by stock analysts.[72]

A comparison of the models (1) to (5) indicates that both the coefficients and the significance levels of the individual characteristics remain virtually unchanged. This finding supports the clear distinction between the individual and the institutional channel. Additionally, the

---

[70] The x-axis of the graphs shows the *BrokerSize* calculated as the log of (1+ number of analysts), 12 analysts corresponds to a log of 2.5 (i.e. (12+1) = e^2.5).

[71] A BrokerSize of 3.7 on the x-axis corresponds with the number of 40 analysts employed at a broker.

[72] The 2.6% higher probability is calculated using the difference in probability of 2 percentage points as indicated on the y-axis in Panel C of Figure 2 on the basis of the public brokers' probability of issuing a bold forecast (77%) at the level of *BrokerSize* of 40 analysts employed; (0.02/0.767)=2.6%.

goodness of fit of the model increases gradually, as documented by the increasing value of the log likelihood chi-square test, explaining that the inclusion of the institutional channel of ownership and underwriting controls yields an improved explanation of the model of analysts' herding behavior. Overall, we find that institutional factors of the brokerage houses have a significant influence on the analysts' CONH, when controlling for a full set of individual analyst characteristics as well as for the broker's underwriting business.

Financial sector analysts vs. analysts covering other sectors

To test the prediction that the effect of the broker's ownership on analysts' forecasting behavior is more pronounced for analysts covering the financial sector, we split the sample into two subsamples according to the two-digit SIC sector codes of the stocks the analysts cover.[73] We define subsample (A) as the *Financial Sector Coverage* and subsample (B) as the *Other Sector Coverage*. We run the logit model from regression equation (2) for each subsample with *Bold* as the dependent variable. We present the results in Table 3. For both subsamples, we start with the reduced models (1) and (5), and we add our main institutional variable of interest *Private* (models (2) and (6)), the interaction term *INT_BrokerSize_Private* (models (3) and (7)), and the *SDC* variables to control for the brokers' underwriting business (models (4) and (8)).[74]

In models (1) and (5), *BrokerSize* is significant in both subsamples, similar to the full sample. In models (2) and (6), the new institutional ownership variable *Private* is not significant in either subsample. However, *BrokerSize* remains significant in both subsamples. We then interact *Private* with *BrokerSize* in models (3) and (7) and find a highly significant and large negative coefficient of *Private* (−0.550) and a highly significant and large positive coefficient of 0.210 for *INT_Private_BrokerSize* for analysts covering the financial sector.

[73] The two-digit SIC codes (from 60 to 69) of the *Financial Sector* defined by the U.S.Department of Labor is identical to Fama and French's industry portfolio number *11. Finance* taken from their compilation of all SIC codes for 12 industry portfolios, available at Kenneth French's homepage.

[74] As we expect the observations to be nested at the brokerage level, we cluster the standard errors at the brokerage level in each of the models. As the brokerage level is superior to the analyst level, the findings are also robust to clustering at the analyst level, which we checked without reporting in Table 3.

Table 3: Regression of Forecast Boldness on Institutional Factors and Analyst Characteristics Split by Analyst Coverage

Dependent variable: *Bold*								
	Column A: Financial Sector Coverage				**Column B: Other Sectors Coverage**			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private		0.0850	-0.5504***	-0.6109***		0.0187	-0.0981	-0.0952
		(1.57)	(-3.67)	(-3.70)		(0.37)	(-0.92)	(-0.97)
INT_Private_BrokerSize			0.2105***	0.2308***			0.0375	0.0384
			(4.17)	(4.15)			(0.91)	(1.07)
BrokerSize	0.0426**	0.0626***	0.0340*	0.0445*	0.0237*	0.0279*	0.0223	0.0000
	(2.23)	(2.84)	(1.68)	(1.84)	(1.85)	(1.88)	(1.42)	(0.00)
LagAccuracy	0.0002	0.0027	0.0017	0.0023	0.1400***	0.1402***	0.1403***	0.1388***
	(0.00)	(0.05)	(0.03)	(0.04)	(6.47)	(6.53)	(6.53)	(6.45)
DaysElapsed	0.0049***	0.0049***	0.0048***	0.0048***	0.0039***	0.0039***	0.0039***	0.0038***
	(5.57)	(5.44)	(5.34)	(5.40)	(9.60)	(9.23)	(9.03)	(9.64)
Horizon	0.0007**	0.0007**	0.0006**	0.0006**	0.0004**	0.0004***	0.0004***	0.0004***
	(2.07)	(2.07)	(2.03)	(2.04)	(2.57)	(2.66)	(2.66)	(2.82)
Frequency	0.0321***	0.0316***	0.0296***	0.0295***	-0.0051	-0.0050	-0.0050	-0.0046
	(2.86)	(2.83)	(2.66)	(2.67)	(-0.96)	(-0.96)	(-0.95)	(-0.88)
GenExperience	0.0113*	0.0117*	0.0118*	0.0119*	0.0028	0.0029	0.0030	0.0023
	(1.76)	(1.82)	(1.85)	(1.86)	(0.86)	(0.88)	(0.91)	(0.73)
FirmExperience	-0.0053	-0.0047	-0.0045	-0.0046	-0.0020	-0.0019	-0.0018	-0.0010
	(-0.76)	(-0.67)	(-0.66)	(-0.68)	(-0.53)	(-0.52)	(-0.48)	(-0.27)
Companies	-0.0035*	-0.0031	-0.0026	-0.0028	-0.0062***	-0.0062***	-0.0063***	-0.0063***
	(-1.81)	(-1.55)	(-1.43)	(-1.42)	(-3.25)	(-3.12)	(-3.16)	(-3.49)
Industries	0.0111	0.0111	0.0090	0.0113	0.0168***	0.0165***	0.0166***	0.0175***
	(0.87)	(0.86)	(0.72)	(0.88)	(3.62)	(3.69)	(3.69)	(4.01)

Table to be continued

Continuation of Table 3

Lead_Underwriter				0.0738				0.1610**
				(0.88)				(2.06)
Co_Underwriter				0.1353				0.1117
				(1.57)				(1.50)
Syndicate_Member				0.1908				0.0830
				(1.53)				(0.89)
Constant	0.7927***	0.6891***	0.8093***	0.6861***	1.0364***	1.0150***	1.0376***	0.9798***
	(6.71)	(4.91)	(6.32)	(4.51)	(13.43)	(12.52)	(12.13)	(10.55)
Clustered by brokerage hous	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,823	21,823	21,823	21,823	114,605	114,605	114,605	114,605
Chi2	41.04	50.14	72.38	70.76	164.3	198.4	229.5	217.2

Notes: The table reports logit coefficient estimates, and in parentheses robust z-statistics clustered by broker. In Column A, we use a subsample of analysts covering firms from the financial sector according to the Standard Industrial Classification (SIC). In Column B, the sample consists of analysts covering all other non-financial sectors. We retain only the last forecast an analyst i is issuing on a specific stock j in year t. *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). *Private* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the natural logarithm of the number of analysts employed at a specific broker. *INT_Private_BrokerSize* is the interaction term between *Private* and *BrokerSize*. We control for the analyst's individual characteristics: *LagAccuracy* is analyst i's last year accuracy on stock j. *DaysElapsed* is the number of days that are elapsed between analyst i's forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst i's forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst i makes during the financial year t for a stock j. *FirmExperience* is the number of consecutive years analyst i covers stock j. *GenExperience* is the number of consecutive years in which analyst i filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst i. In models (4) and (8), we include controls for underwriting activities: *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year t as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia and company home pages.

The pattern is similar to that in the full sample; however, the magnitude of the coefficients of *Private* and *INT_Private_BrokerSize* in subsample (A) is about three times larger than in the full sample, indicating a strongly pronounced effect for financial sector analysts compared to subsample (B). Particularly in the subsample of analysts covering other sectors (B), neither the coefficient for *Private* nor the coefficient for *INT_Private_BokerSize* is significant. This pattern remains the same when we control for the full set of underwriting variables in model models (4) and (8), respectively. For financial sector analysts, the coefficients of −0.61 and 0.23 for *Private* and *INT_Private_BrokerSize*, respectively, are both highly significant and substantially larger in magnitude than the coefficients for analysts covering other sectors (i.e., −0.09 and 0.04 for *Private* and *INT_Private_BrokerSize*, respectively). Both differences between the two subsamples, for *Private* (−0.52) and for *INT_Private_BrokerSize* (0.19), are significant at the 1% level, as shown by a two-sided *t*-test run with two seemingly unrelated regressions (untabulated). The results confirm that the effect of the broker's ownership is considerably more pronounced for financial analysts than for analysts covering other sectors. As only financial sector analysts can exert a negative externality on their employer's valuation, they face higher costs when producing bold forecasts than analysts covering other sectors if working at a publicly listed brokerage.

We further analyze these findings and examine the marginal effects of subsample (A), *Financial Sector Coverage*. We compute the AME between working with a privately held or a publicly listed brokerage (*Private*), yielding a positive and highly significant difference in the probability of issuing a bold forecast for analysts at private brokers of 4.8% at the 1% significance level.[75] When compared to an individual characteristic such as general experience (*GenExperience*), analysts would have to gain additional experience of about 18 years to have the same effect on their probability of issuing a bold forecast.[76] For subsample (B), *Other Sectors Coverage*, the difference in AME between analysts at private and public brokerage houses, is much lower (1.2%) and insignificant (*z*-statistic = 0.89).

Similar to the full sample, we further investigate the marginal effects of ownership by plotting AME conditional on *BrokerSize* for both types of ownership over the common support of

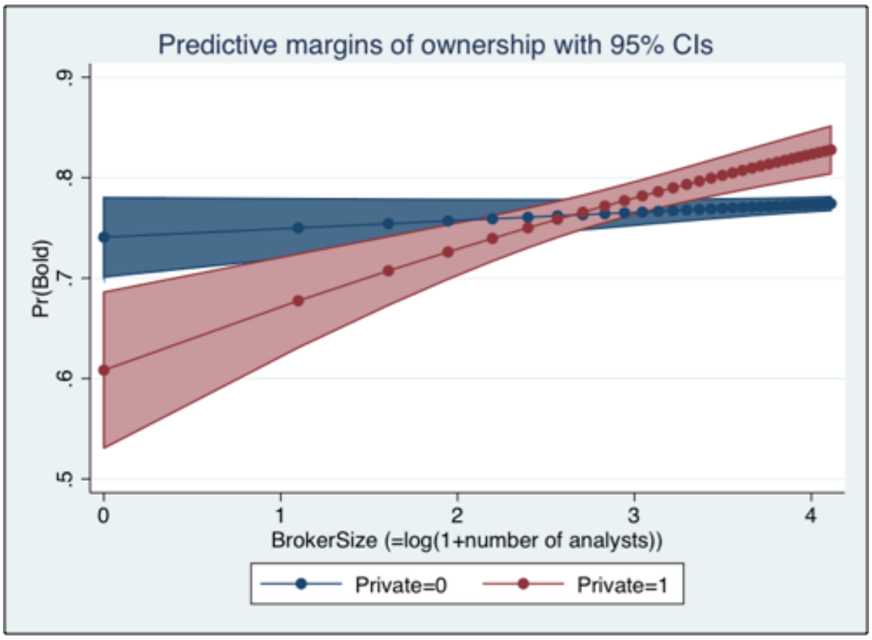
[75] Likewise, the DEP at means shows a difference in probability of 5.3% at the 1% significance level.

[76] The comparison of the two different effects is calculated by dividing the AME of Private by the AME of GenExperience, yielding a ratio of 18.17 (years).

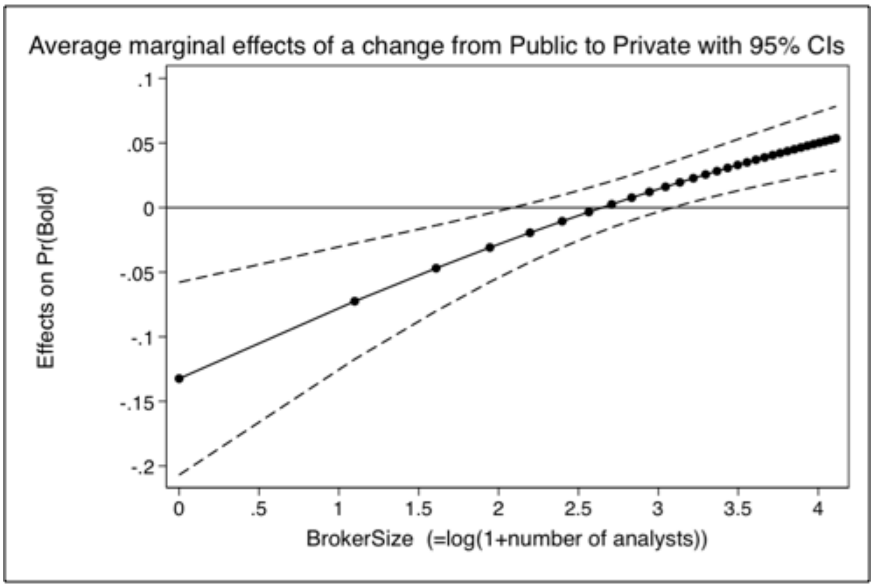
BrokerSize, which ranges from 0 to 61 analysts.[77] The plots for subsample (A) and (B) are shown in Figure 3 and Figure 4, respectively.

Figure 3: Plotted Marginal Effects of *Private* for Different *BrokerSizes* of Subsample (A) Financial Sector Coverage

Panel A



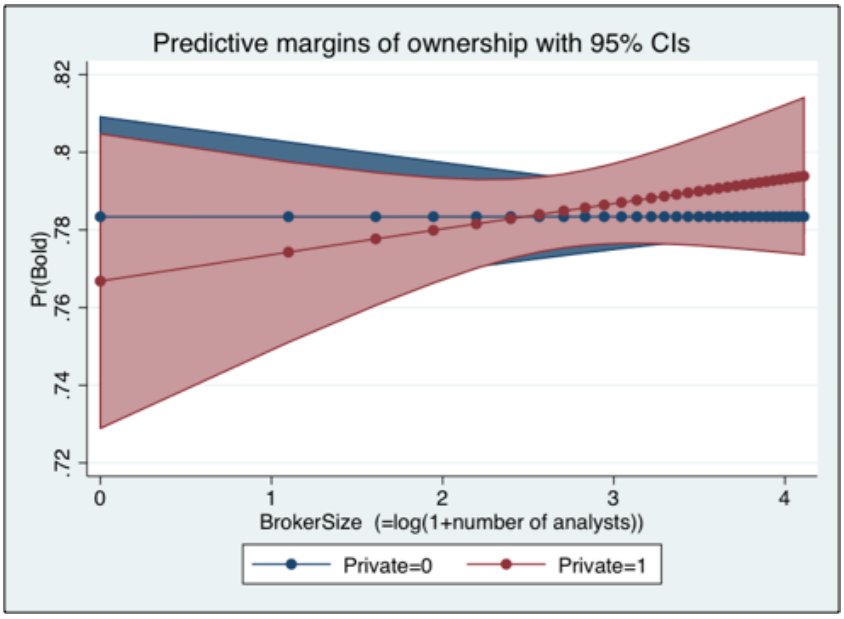
Panel B



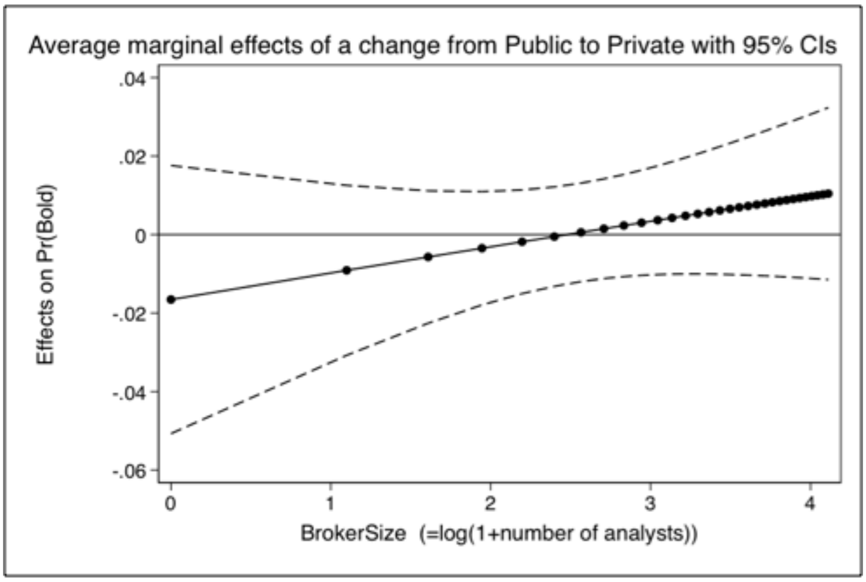
[77] As with the full sample, we determine the underwriter level to be the highest (i.e., Lead_underwriter = 1), because the majority of the observations in both subsamples also belong to this group. All other underwriting variables (Co_Underwriter, Syndicate_Members and No_Underwriting) are kept at zero.

Figure 4: Plotted Marginal Effects of Private for Different BrokerSizes of Subsample (B) Other Sectors Coverage

Panel A



Panel B



The plots confirm the earlier findings. Panel A in Figure 3 shows that the marginal effect of ownership increases with the size of the brokerage house. Panel B reveals that financial sector analysts working for privately held houses that employ more than 24 analysts (i.e. *BrokerSize* equals 3.2 in the figure) have a significant difference at the 5% level in the probability of issuing a bold forecast when compared to financial sector analysts at public brokers. At this level of *BrokerSize*, the difference reflects a 3% higher probability for financial sector analysts from private brokers than analysts from public brokers to release a bold forecast.[78] Because the difference increases with the number of analysts employed, financial sector analysts from private brokers with more than 60 analysts have a 6.5% higher probability of issuing a bold forecast than similar analysts from public brokers.[79] This accounts for an economically significant difference in analysts' herding behavior. To reach the same effect with an individual characteristic, for example, analysts with 1 year of general experience would have to gain more than 27 years of additional experience.[80] In contrast, Panel B of Figure 4 shows that, for analysts covering other sectors (subsample B), the difference between analysts from private brokers and analysts from listed brokers is both insignificant and substantially smaller in magnitude than for financial sector analysts for each level of *BrokerSize*, except around the break-even point.

Overall, we provide evidence that the ownership effect on the analysts' CONH is driven by the group of analysts covering the financial sector. As financial sector analysts working with publicly listed brokerage houses exert externalities on the valuation of their own employer, they can forecast less independently. They face higher CONH and, hence, are more prone to herding behavior and issue fewer bold forecast revisions than analysts who work at privately held brokers, which are not dependent on the stock market. In contrast, for analysts covering other sectors, the ownership effect vanishes. Our analysis provides evidence that the ownership effect does not generally lead to a difference in forecasting behavior for all analysts; instead, it is the group of analysts covering the financial sector that drives the difference.

It is important to emphasize that, similar to the full sample, the coefficients of the individual analyst variables are essentially uninfluenced by the inclusion of the institutional variables for ownership (*Private*) and the variables for the brokers' underwriting business. In addition, the

[78] As before, the 3% higher probability is calculated using the difference in probability of 2.3 percentage points as indicated on the y-axis in Panel B of Figure 3 on the basis of the public brokers' probability of issuing a bold forecast (77%) at the level of BrokerSize of 24 analysts employed; (0.023/0.767)=3.0%.

[79] The 6.5% higher probability is calculated as above, but at the level of 60 analysts employed; (0.05/0.774)=6.5%.

[80] We calculate the comparison of the marginal effect by dividing the AME of *Private* at a *BrokerSize* level of 60 analysts employed by the AME of one year additional experience (*GenExperience)*; this yields a ratio of 27.67 (years).

ownership effect of private brokers is not affected when we control for underwriting business. This result increases confidence in our tests.[81]

Although we do not focus on individual analyst characteristics, using them instead as controls, we note two significant differences of individual characteristics between analysts covering the financial sector and analysts covering other sectors. The effect of lagged accuracy on the probability of issuing a bold forecast essentially does not exist for financial sector analysts but is significantly positive for the other sectors in subsample (B) (difference = −0.137, significant on the 2% level, two-sided *t*-test, untabulated). Scharfstein and Stein (1990), Trueman (1994) and Clement and Tse (2005) show theoretically and empirically that the herding behavior of analysts depends on their confidence in their own abilities. The absence of the influence of self-assessed ability, measured by lagged accuracy, on the probability of issuing a bold forecast (in all models (1) to (4) of subsample A) highlights another difference in CONH between financial sector analysts and other sector analysts. The finding suggests that financial sector analysts do not rely on their self-assessed ability when revising their forecasts.

In contrast, the significant difference of *Frequency* between subsample (A) and (B) is an indicator of the analysts' independent ability to reveal new information to the market in a timely manner. *Frequency* is only significant and positive for the financial sector subsample and the difference between subsamples (A) and (B) is highly significant at the 1% level (two-sided *t*-test, untabulated). To the extent that financial sector analysts seem to face higher CONH if working for public brokers, they might be discouraged from frequently revising their forecasts. Thus, *Frequency* bears an additional characteristic to identify more or less independent financial sector analysts.

4.5.3. Robustness Checks

We perform three robustness checks. First, we restrict the sample to observations that share a common support in *BrokerSize*, which lies between 1 and 61 analysts employed, defined by the largest privately-held brokerage house that employs a maximum of 61 analysts. Due to the restriction, we lose about 35% of the observations of the full sample. We then run the logit model from regression equation (2) for both subsamples – (A), *Financial Sector Coverage*, and (B), *Other Sectors Coverage*. The coefficients of the two models (1) and (2) presented in

[81] In addition, the results hold when running various robustness checks (see next section).

Table 4 exhibit the same pattern of coefficient size and significance as in the unrestricted sample (i.e., models (4) and (8) of Table 3). The coefficients of *Private* and *INT_Private_BrokerSize* of –0.55 and 0.21, respectively, remain highly significant (1% level) with similar magnitude for the financial sector sample, whereas for the other sectors sample the coefficients remain insignificant and are of lower magnitude. Thus, the results with a common support sample confirm the ownership effect and its considerably different impact on analysts covering financial sector stocks.

Secondly, we include year-stock fixed effects in the original model in regression equation (2). In a model analyzing the bias of analyst forecast errors, Clement (1999) shows that controlling for year-stock effects (1) raises the probability of identifying systematic differences in analysts' forecast errors, and (2) reduces possible heteroskedasticity due to larger variations in the forecast errors for stocks with larger earnings per share. Thus, although Clement and Tse (2005) do not use year-stock fixed effects in their analysts' herding behavior model, we apply the year-stock as a robustness check. Models (3) and (4) unveil the identical pattern of magnitude and significance for the coefficients of *Private* and the interaction term between *Private* and *BrokerSize*. This again supports the finding that the analysts covering financial sector stocks drive the effect of the broker's ownership on the analysts' CONH and, hence, on their herding behavior.

In our setting, the use of analyst fixed effects to control for unobservable differences in individual analyst characteristics is not feasible. Analyst fixed effects eliminate the constant institutional effects of the analysts' brokerage house, such as ownership. We would "throw out the baby with the bathwater" when using an analyst fixed effects model for our analysis of constant institutional brokerage characteristics. Furthermore, as we employ a binary dependent variable, the use of analyst fixed effects drops all the observations of analysts consistently issuing bold or herding forecasts over time (e.g., about one third of the subsample of analysts covering the financial sector). However, we can employ an analysts' random-effects model, which uses not only the within-group but also the between-group variation and, hence, does not drop any observations due to invariability. We obtain highly significant coefficients in model (3) in Table 4 of the identical pattern as shown before in Table 3, model (4): *Private* yields a negative sign, *INT_Private_BrokerSize* yields a positive sign, and both coefficients remain at nearly the identical magnitude as in the original model.

Table 4: Regression of Forecast Boldness on Institutional Factors and Analyst Characteristics: Robustness Checks

Dependent variable: *Bold*	**Common Support**		**Year-Stock FE**		**OLS**	
	Financial Sector	Other Sectors	Financial Sector	Other Sectors	Financial Sector	Other Sectors
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Private	-0.5475***	0.0035	-0.6395***	-0.0273	-0.1108***	-0.0162
	(-2.60)	(0.03)	(-2.86)	(-0.28)	(-3.55)	(-0.94)
INT_Private_BrokerSize	0.2088***	0.0092	0.2300***	0.0013	0.0413***	0.0065
	(3.05)	(0.24)	(3.19)	(0.04)	(4.00)	(1.05)
BrokerSize	0.0552	0.0314	0.0354	0.0070	0.0078*	0.0001
	(1.17)	(1.23)	(1.20)	(0.52)	(1.81)	(0.02)
LagAccuracy	-0.0148	0.1202***	0.0285	0.1477***	0.0005	0.0239***
	(-0.25)	(3.85)	(0.39)	(4.52)	(0.05)	(6.33)
DaysElapsed	0.0060***	0.0038***	0.0055***	0.0066***	0.0008***	0.0006***
	(5.78)	(7.22)	(4.25)	(11.20)	(5.64)	(9.80)
Horizon	0.0011***	0.0007***	0.0021***	0.0013***	0.0001**	0.0001***
	(2.86)	(3.60)	(5.23)	(7.85)	(2.04)	(2.78)
Frequency	0.0531***	0.0092	0.0088	-0.0067	0.0051***	-0.0008
	(4.14)	(1.38)	(0.70)	(-1.44)	(2.72)	(-0.88)
GenExperience	0.0100	-0.0032	0.0170**	0.0015	0.0020*	0.0004
	(1.12)	(-0.87)	(2.19)	(0.47)	(1.89)	(0.73)
FirmExperience	-0.0059	0.0036	-0.0119	-0.0019	-0.0008	-0.0002
	(-0.65)	(0.77)	(-1.18)	(-0.45)	(-0.67)	(-0.28)
Companies	-0.0024	-0.0091***	-0.0006	-0.0016	-0.0005	-0.0011***
	(-1.10)	(-3.70)	(-0.20)	(-1.23)	(-1.40)	(-3.30)
Industries	0.0062	0.0176***	-0.0115	-0.0097*	0.0020	0.0030***
	(0.43)	(3.84)	(-0.67)	(-1.85)	(0.89)	(4.13)

To be continued

Continuation of Table 4

Lead_Underwriter	0.0737	0.1594**	0.1254*	0.2140***	0.0134	0.0284**
	(0.94)	(2.18)	(1.75)	(6.92)	(0.87)	(1.98)
Co_Underwriter	0.1444*	0.1141*	0.1291	0.2146***	0.0242	0.0199
	(1.79)	(1.68)	(1.44)	(5.62)	(1.54)	(1.45)
Syndicate_Member	0.1898	0.0872	0.1996	0.0589	0.0331	0.0149
	(1.56)	(1.02)	(1.57)	(1.04)	(1.46)	(0.89)
Constant	0.5205**	0.8633***	–	–	0.6784***	0.7306***
	(2.28)	(8.03)	–	–	(24.93)	(44.24)
Clustered by brokerage house	Yes	Yes	Yes	Yes	Yes	Yes
Year-stock FE	No	No	Yes	Yes	No	No
Observations	14,756	67,531	15,033	80,961	21,823	114,605
R-squared	–	–	–	–	0.003	0.002
Number of year-stock	–	–	2,089	10,459	–	–
Chi2	74.31	141.7	78.57	337.9	–	–

Notes: The table reports regression results for both analyts covering the financial sector (models 1, 3, and 5) and analysts covering other sectors (models 2, 4, and 6). In models (1) and (2), we use a common support sample by excluding brokers with more than 61 analysts (i.e., the maximum broker size of private brokers, see Table 1). In models (3) and (4), we include year-stock fixed effects. Models (5) and (6) report OLS coefficient estimates, and in parentheses, robust t-statistics clustered by broker. Models (1) to (4) report logit coefficient estimates, and in parentheses, robust z-statistics clustered by broker. We retain only the last forecast an analyst i is issuing on a specific stock j in year t. The dependent variable *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). *Private* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the natural logarithm of the number of analysts employed at a specific broker. *INT_Private_BrokerSize* is the interaction term between *Private* and *BrokerSize*. We control for the analyst's individual characteristics: *LagAccuracy* is analyst i's last year accuracy on stock j. *DaysElapsed* is the number of days that are elapsed between analyst i's forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst i's forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst i makes during the financial year t for a stock j. *FirmExperience* is the number of consecutive years analyst i covers stock j. *GenExperience* is the number of consecutive years in which analyst i filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst i. We also include controls for underwriting activities in all models: *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year t as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia, and company home pages.

A *Hausman* test comparing the analysts' random effects model and the analysts' fixed effects model shows that the difference in coefficients is marginally unsystematic (p-value of 0.104), providing further robustness of the ownership effect.

Thirdly, to once more tackle the concerns regarding the interpretation of interaction terms in non-linear models, we perform our tests by using a linear (OLS) regression model. Table 4 models (5) and (6) show the OLS output for subsamples (A) and (B), revealing the identical pattern of coefficient magnitudes and significance levels as before in the logit models (4) and (8) in Table 3. The coefficients of *Private* and *INT_Private_BrokerSize* in model (5) of Table 4 are both highly significant (both robust |t-statistics| >3.5) and have a negative and a positive sign, respectively. For analysts of model (6) in Table 4, *Other Sectors Coverage*, the same variables show a much lower coefficient (almost factor ten) and no significance, which, again, is identical to the findings from the logit model (8) in Table 3.

4.5.4. Accuracy, Boldness, and the Ownership Effect

So far, we have shown that, firstly, financial sector analysts working with privately held brokerage houses have lower CONH and, thus, a higher probability of issuing bold forecasts, and that, secondly, this effect is primarily driven by the subsample of analysts covering the financial sector. Now, the question arises whether the bold forecasts actually contain more relevant information, in other words, whether bold forecasts are more accurate. Clement and Tse (2005) have shown that bold forecasts have a higher forecast accuracy compared to herding forecasts. However, we also test for our sample whether (1) analysts who issue bold forecasts rely on relevant private information and, thus, issue more accurate forecasts on average, or (2) whether they are "overconfident but poorly informed" (Clement and Tse 2005, p. 317). In the latter case, there would be no association between boldness and the analysts' accuracy. We do not predict a direct effect of ownership on forecast accuracy; we expect that the broker's ownership indirectly affects accuracy through a higher probability of issuing bold forecasts, as shown in the previous section.

As before, we explore the difference between analysts that cover the financial sector and analysts that cover other sectors. We thus split the sample into subsamples (A), *Financial Coverage,* and (B), *Other Sectors Coverage*. Similar to Clement and Tse (2005), we regress our measure of forecast accuracy (*Accuracy*) on forecast boldness (*Bold*), our measure of ownership (*Private*), the broker's size (*BrokerSize*), the interaction term

(INT_*Private*_BrokerSize), and the individual and institutional controls. The equation of the linear (OLS) regression is as follows:

$$\text{Accuracy} = \alpha_0 + \alpha_1 \text{ Bold}_{ijt} + \alpha_2 \text{ Private}_{ijt} + \alpha_3 \text{ BrokerSize}_{ijt}$$
$$+ \alpha_4 \text{ INT\_BrokerSize\_Private}_{ijt} + \alpha_m \text{ SDC}^m_{ijt} + \alpha_k X^k_{ijt} + \varepsilon_{ijt} \quad , \qquad (3)$$

where *Accuracy* is as defined in equation (1) and all the other variables are as defined in regression equation (2).

Table 5 shows the OLS results where *Accuracy* is the dependent variable. Models (1), (2), (5), and (6) do not have year-stock fixed effects, whereas models (3), (4), (7), and (8) include year-stock fixed effects as an additional refinement of our model.[82] To provide a consistent setting throughout the paper, the first models of each pair are regressed without *Bold*, and have the same set of independent variables as in the earlier regression equation (2), where *Bold* was the dependent variable. We are thus able to compare the influence of the addition of *Bold* on the other relevant variables. The second models of each pair are estimated as shown in regression equation (3).

The results from models (1) and (5) of Table 5 show that the ownership has no direct effect on the accuracy of analysts' forecasts in both subsamples. The coefficients *Private* and *INT_Private_BrokerSize* have no statistical significance, but have the same sign as before, when *Bold* was the dependent variable. When adding *Bold* in models (2) and (6) as independent variable, the variable is highly significant at the 1% level and positive. In both subsamples, we confirm the results of Clement and Tse (2005) that bold forecasts convey new relevant private information that was not yet included in other analysts' forecasts. Accordingly, *ownership* has only an indirect effect on forecast accuracy through the increase in likelihood of issuing bold forecasts for analysts from private brokers above a critical size (*BrokerSize*).

Interestingly, if we control for year-stock effects in models (3) and (4), and (7) and (8), the direct influence of the brokerage's ownership becomes significant for subsample (A) *Financial Sector Coverage*: Privately held brokerage houses do marginally exhibit a direct influence on the forecast accuracy of their financial sector analysts.

[82] See also Clement (1999). Analyzing analyst forecast errors, he shows that controlling for year-stock effects, firstly, raises the probability of identifying systematic differences in analysts' forecast errors, and secondly, reduces possible heteroskedasticity due to larger variations in the forecast errors for stocks with larger earnings per share.

Table 5: Regression of Forecast Accuracy on Forecast Boldness, Institutional Factors and Analyst Characteristics

Dependent variable: ***Accuracy***								
	Column A: Financial Sector Coverage				**Column B: Other Sectors Coverage**			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bold		0.0482***		0.0548***		0.0456***		0.0500***
		(10.09)		(10.18)		(18.21)		(22.05)
Private	-0.0533	-0.0477	-0.0638**	-0.0579**	0.0007	0.0015	-0.0066	-0.0063
	(-1.44)	(-1.30)	(-2.54)	(-2.31)	(0.02)	(0.05)	(-0.63)	(-0.61)
INT_Private_BrokerSize	0.0082	0.0062	0.0148*	0.0127	-0.0051	-0.0054	0.0019	0.0018
	(0.65)	(0.49)	(1.84)	(1.59)	(-0.45)	(-0.48)	(0.58)	(0.57)
BrokerSize	0.0024	0.0020	-0.0054*	-0.0056*	-0.0003	-0.0003	-0.0036**	-0.0036**
	(0.45)	(0.41)	(-1.68)	(-1.77)	(-0.12)	(-0.12)	(-2.52)	(-2.57)
LagAccuracy	0.0447***	0.0448***	0.0297***	0.0295***	0.0721***	0.0710***	0.0444***	0.0433***
	(5.66)	(5.69)	(3.73)	(3.72)	(13.43)	(13.37)	(12.69)	(12.39)
DaysElapsed	-0.0008***	-0.0008***	0.0000	-0.0000	-0.0007***	-0.0008***	-0.0001	-0.0001**
	(-4.93)	(-5.30)	(0.22)	(-0.12)	(-16.71)	(-17.32)	(-1.46)	(-2.27)
Horizon	-0.0014***	-0.0014***	-0.0017***	-0.0017***	-0.0013***	-0.0013***	-0.0015***	-0.0015***
	(-25.94)	(-26.58)	(-38.59)	(-39.07)	(-48.90)	(-49.26)	(-84.68)	(-85.47)
Frequency	0.0003	0.0001	0.0027**	0.0027*	0.0027***	0.0027***	0.0019***	0.0020***
	(0.28)	(0.09)	(1.96)	(1.93)	(5.59)	(5.59)	(3.75)	(3.87)
GenExperience	-0.0011	-0.0012	0.0003	0.0001	-0.0015***	-0.0015***	0.0006	0.0005
	(-0.97)	(-1.07)	(0.34)	(0.17)	(-3.72)	(-3.79)	(1.61)	(1.59)
FirmExperience	0.0030**	0.0030**	-0.0002	-0.0001	0.0033***	0.0033***	-0.0000	-0.0000
	(2.35)	(2.41)	(-0.15)	(-0.05)	(6.43)	(6.48)	(-0.11)	(-0.08)
Companies	-0.0003	-0.0003	0.0006*	0.0006*	0.0004	0.0004	-0.0002	-0.0002
	(-0.95)	(-0.88)	(1.89)	(1.91)	(1.32)	(1.47)	(-1.43)	(-1.33)
Industries	0.0004	0.0002	-0.0041**	-0.0040**	-0.0058***	-0.0059***	-0.0004	-0.0004
	(0.19)	(0.13)	(-2.22)	(-2.17)	(-7.76)	(-8.03)	(-0.79)	(-0.68)

To be continued

Continuation of Table 5

Lead_Underwriter	0.0058	0.0052	0.0288***	0.0277***	0.0094	0.008	0.0246***	0.0228***
	(0.83)	(0.74)	(3.60)	(3.48)	(1.18)	(1.02)	(7.39)	(6.87)
Co_Underwriter	-0.0016	-0.0026	0.0058	0.0048	0.0184**	0.0174**	0.0228***	0.0209***
	(-0.10)	(-0.16)	(0.59)	(0.49)	(2.27)	(2.18)	(5.53)	(5.10)
Syndicate_Member	0.0136	0.0122	0.0044	0.0027	-0.0097	-0.0104	-0.0020	-0.0026
	(0.94)	(0.85)	(0.31)	(0.19)	(-0.61)	(-0.65)	(-0.34)	(-0.44)
Constant	0.7862***	0.7530***	0.8110***	0.7732***	0.7776***	0.7445***	0.7986***	0.7631***
	(29.26)	(27.87)	(41.41)	(38.91)	(65.27)	(62.02)	(103.14)	(96.75)
Clustered by brokerage hous	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-stock FE	No	No	Yes	Yes	No	No	Yes	Yes
Observations	21,041	21,041	21,041	21,041	108,588	108,588	108,588	108,588
Number of year-stock	–	–	4,166	4,166	–	–	19,451	19,451
R-squared	0.073	0.077	0.111	0.116	0.077	0.081	0.102	0.107

Notes: The table reports OLS regression coefficients, and in parentheses, robust *t*-statistics clustered by broker. The dependent variable *Accuracy* is the standardized accuracy of analyst *i* covering firm *j* in the year *t*. *Accuracy* is calculated as the maximum absolut forecast error (AFE) for analysts following firm *j* in year *t* minus the AFE for analyst *i* following firm *j* in year *t,* scaled by the range (max - min) of AFE for analyts following firm *j* in year *t*. The main independent variable *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). We include *Bold* in models (2), (4), (6), and (8), and we include year-stock fixed effects in models (3), (4), (7), and (8). In Column A, we use a subsample of analysts covering firms from the financial sector according to the Standard Industrial Classification (SIC). In Column B, the sample consists of analysts covering all other non-financial sectors. We retain only the last forecast an analyst *i* is issuing on a specific stock *j* in year *t*. *BoldPrivate* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the natural logarithm of the number of analysts employed at a specific broker. *INT_Private_BrokerSize* is the interaction term between *Private* and *BrokerSize*. We control for the analyst's individual characteristics: *LagAccuracy* is analyst *i*'s last year accuracy on stock *j*. *DaysElapsed* is the number of days that are elapsed between analyst *i*'s forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst *i*'s forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst *i* makes during the financial year *t* for a stock *j*. *FirmExperience* is the number of consecutive years analyst *i* covers stock *j*. *GenExperience* is the number of consecutive years in which analyst *i* filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst *i*. We include controls for underwriting activities in all models: *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year *t* as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia, and company home pages.

In model (3), the coefficient of *Private* is significantly negative (−0.06; robust *t*-statistic: −2,54) and the coefficient of the interaction term *INT_Private_BrokerSize* is positive and weakly significant (0.01; robust *t*-statistic: 1.84). The results indicate that financial sector analysts working with a privately held brokerage that employs more than five analysts have a higher accuracy than analysts with listed brokers. The results virtually remain the same when controlling for *Bold* in model (4). Similar to the results when *Bold* is the dependent variable, we do not find significant evidence of an ownership effect for analysts covering other sectors (models (5) – (8)).

Thus, in both subsamples, we confirm the results of Clement and Tse (2005) that bold forecasts convey new relevant private information that was not yet included in other analysts' forecasts. Accordingly, ownership bears a mostly indirect effect on forecast accuracy through the reduction in CONH and the consequent increase in likelihood of issuing bold forecasts for analysts from private brokers with sufficient resources (i.e., number of analysts employed).

When controlling for year-stock effects, the coefficient of *BrokerSize* becomes significantly negative in both subsamples, indicating that analysts from public brokers with more resources and/or higher reputation do not necessarily produce more accurate forecasts, but rather more inaccurate ones. Finally, it is notable that the self-assessed ability (measured with *LagAccuracy*) in each of the models is highly significant and substantially positive. Thus, self-assessed ability is clearly one of the drivers of accurate forecasts, as it is a necessary condition for an analyst to have confidence in his or her private information.

4.6. Conclusion

As security analysts play a key role in providing information to capital markets, analysts' herding behavior results in a loss of information and leads to a misallocation of scarce resources. To avoid this waste of resources, it is important to understand the drivers of analysts' herding behavior. Prior literature has focused on *individual* analyst characteristics and herding behavior (Hong et al. 2000; Clement and Tse 2005; Bernhardt et al. 2006; Jegadeesh and Kim 2010). The literature on analysts' herding behavior is largely silent about the influence of the *institutional* environment. If any, broker size as a proxy for reputation or available resources is the only institutional factor that appears in previous research on analysts' herding behavior. Only the related literature on analysts' forecast errors explicitly addresses the brokers' underwriting business as an institutional influence on forecasting behavior and touches on other determinants such as geographical distance.

This study examines the influence of institutional factors on the herding behavior of security analysts. For this purpose, we provide a general framework combining individual and institutional factors. The framework allows the analysis of herding behavior based on the costs of non-herding (CONH). High CONH indicate that a specific analyst has to bear high psychological, social, or economic costs if he or she does not follow the prevailing opinion (i.e. the herd), and that such analysts are thus more prone to herding behavior than other analysts facing low CONH.

Based on the framework, we introduce and examine the influence of a new institutional factor, ownership, on the herding behavior of security analysts. Specifically, we examine, firstly, whether analysts working at publicly listed brokerage houses are more likely to herd than analysts from privately held brokers, and secondly, if this effect is different for analysts covering their own sector (i.e., for financial sector analysts). According to Stein's (2002) theory of soft and hard information production in firms, we predict that analysts from publicly listed brokers are less independent and, hence, more prone to herding behavior than analysts from private brokers. Analysts from public brokers face higher CONH (e.g., effort, scrutiny, or longer chains of command) when processing bold information. The processing of information that deviates from a prevailing opinion (i.e. the herd) is more difficult in more centralized and regulated environments with taller hierarchies, holding size and scope fixed (Stein 2002 and Berger et al. 2005). In contrast, privately held brokerage houses share generally flatter hierarchies and more decentralized operations than public brokers.

To test the effect of the broker's ownership, we hand-collected data on whether U.S. brokers are publicly listed or privately held. We enriched this unique data set with data about the brokers' underwriting business and their informational environment and included a full set of controls for the individual analyst characteristics.

We find that the broker's ownership (i.e., public vs. private) has a substantial effect on its analysts' CONH and thus on their herding behavior when issuing forecasts. Second, concerning the analysts' CONH, we reveal a tradeoff between the broker's size and the broker's ownership. Lack of resources and a lower reputation seem to constitute handicaps for analysts working at small (and, hence, mostly private) brokerages, making them more prone to herding behavior. However, the effect that private brokers process soft information more easily, resulting in lower CONH for analysts when issuing a bold forecast, begins to dominate after a threshold of 40 analysts is reached. Analysts working with more than 40 colleagues at a privately held brokerage exhibit a significantly higher probability of issuing a bold (i.e. non-herding) forecasts of 2.6% than similar analysts working at public brokers.

Third, we provide strong evidence that this effect is driven by the group of analysts covering the financial sector; financial sector analysts at publicly listed brokers face substantially higher CONH and are, hence, more prone to herding behavior than analysts at private brokers. We argue that this is due to the externalities that financial sector analysts working at publicly listed brokers can exert on the valuation of their employer and on related issues such as their executives' compensation. Financial sector analysts working at privately held brokerage houses employing 24 or more analysts show significantly lower CONH than when employed at a publicly listed brokerage. At this threshold, the probability of issuing a bold (i.e. non-herding) forecast is on average 3.0% higher for analysts from larger private brokers than for analysts from listed brokers. As the difference increases with the number of analysts employed, the probability of issuing a bold forecast for analysts from private brokers with more than 60 analysts is 6.5% higher than for the same analysts from a publicly listed broker, which accounts for an economically meaningful difference in herding behavior.

Fourth, we confirm the findings by Clement and Tse (2005) that these bold forecasts (i.e. non-herding forecasts) are more accurate, and not simply the result of overconfident or poorly informed analysts. Fifth, the effect is robust to a battery of additional checks, all supporting our results.

Taken together, our insights provide strong evidence that the institutional factors of brokerage houses substantially shape the herding behavior of their analysts. We show that the ownership effect is an important new institutional factor in our framework. The effect constitutes a different channel from earlier institutional determinants of analysts' forecasting behavior, such as the volume of investment banking, the reputation, or the resources of a brokerage house.

For investors and regulators, our findings indicate that information about financial sector stocks is more independent, and thus, less prone to herding behavior, if it stems from analysts working at privately held brokerage houses that are large enough to provide their analysts with adequate resources.

Our study also has implications for resolving the current reliability issues of information dissemination in financial markets. The literature mainly cites two sources for fundamental firm information, financial analysts and credit rating agencies. There is evidence that these two players discipline each other's actions (Fong et al. 2012). However, for financial sector stocks, an agency problem seems to occur for both players. First, rating agencies are paid by financial firms to provide ratings for their financial products, and hence, the agencies are often reluctant to reveal bold information about their clients, as has been learned in the aftermath of the last financial crisis (Mathis et al. 2009). Second, our findings suggest that financial sector analysts from publicly listed brokerages have incentives to herd and not to reveal bold information about the financial stocks they cover. So, taken together, the two most important players in the market for information about financial sector stocks (i.e., the rating agencies and the analysts from the large, publicly listed brokerage houses) seem to have only limited incentives to reveal genuine private information about financial sector stocks. Thus, it is important to shed further light on the institutional characteristics of information providers in financial markets and on how the institutional factors can be changed. These changes should aim at increasing the amount of independent information available about the future condition of financial sector companies. This may result in a more efficient asset and resource allocation in financial markets and could increase the probability of detecting a looming financial crisis earlier.

5. Forecasts and Reactivity: A Theoretical Examination[83]

5.1. Introduction

Forecasts are ubiquitous. There is virtually no place in society from which forecasts are excluded. Before we decide to take the train, an application on our mobile phone tells us which train to choose in order to avoid the bulk of passengers, or, for car drivers, which route will be the fastest. The future freshman tries to forecast the best college place based on the latest undergraduate school rankings. Households try to predict periods of particularly low real estate prices when searching for a new house. Firms seek to predict the best moment to invest in a new product, industry or country. Shareholders try to predict the highest prices to sell their shares, and governments issue predictions about the general economic outlook and take measures according to it.

But there is always one fundamental question: Which forecast should we follow and how much do we believe in it? Maybe we follow the forecast provider that issued the most accurate forecast last time and assume that the same provider will also be successful next time. But what if most of the train passengers employ the same forecast service, because it was the most accurate one in the previous period, and therefore switch to the same train? The train will be crowded. This time, the formerly best forecast service might no longer be perceived as the most accurate one although it had been ex-ante (before the forecast was publicly issued). In contrast, what if all the talented students apply to the university that is highest ranked in the recent college ranking? This university will acquire the most talented applicants and, thus, the best selection of students. It will therefore be ranked number one in the future rankings, although ex-ante the ranking might have been wrong.

In markets in which participants react to forecasts, the ex-post analysis of the forecast accuracy is always doubtful. Given a high degree of reactivity in a market, individuals' reliance on the previous periods' best-performing forecasters can create a self-fulfilling prophecy. The accuracy of a forecaster's past predictions provides no guarantee that future

[83] This chapter is based on a paper written jointly with Bruno S. Frey (Cueni and Frey 2013).

predictions will be correct. However, the fact that many individuals believe the prediction of the best former forecaster to be the most correct one and act upon it makes the market move in the direction of this prediction. Under such circumstances, it becomes rational to follow the herd and to stick to the forecaster with the most accurate past prediction, because it will become the most correct one again due to the self-fulfilling prophecy induced by reactivity. In the train example, in contrast, a self-defeating prophecy may arise if a large majority of people starts following the previously most accurate forecast provider.

Reactivity is not only of concern for the forecast users (i.e. market participants). The forecasters themselves suffer equally from reactivity. It affects the actual outcome and thus hampers the ex-post calibration of their models and their learning in general. Reactivity in a market makes the distinction between a good and a bad forecast very difficult or in many cases impossible. Although the forecasters and the market participants may act in a fully rational way, they are both deceived by reactivity.

Of course, there are early contributors to the analysis of reactive behavior in the social sciences. Venn (1866), Morgenstern (1928) and Merton (1936; 1948) approached the problem of reactivity in anthropogenic markets using concepts such as self-fulfilling or self-defeating prophecies.[84] This paper pursues four aims. First, we intend to draw attention once more to the forecasting problems that are evoked by reactivity and that cannot be solved with more data or faster information technology. The modern economic forecasting literature is mostly enthusiastic about the amount of new data (big data) that has become accessible due to the new information technologies. It points toward an increasing ability to provide much more accurate and timely forecasts in the near future (see, e.g., Elliott and Timmermann 2008; Clements and Hendry 2011, Part II and III). We intend to show that the new technological possibilities do not improve forecasts in reactive markets, and we support our proposition with empirical evidence.

Second, we point out several areas where the neglect of reactivity by scientists may lead to biased results. Examples are security analysts' forecasts (see, e.g., Kothari 2001; Clement and Tse 2005; Beyer 2008; Liu and Natarajan 2012), art auctioneers' presale price forecasts (see,

[84] Venn (1866, p. 187) used the term "suicidal prophecies", similar to Morgenstern's (1928, p. 107) "[...]the publicized prediction has destroyed its own foundation." (Translated into English by the authors, but stated originally in German: *"[...] die bekanntgegebene Prognose [hat] ihre eigene Basis zerstört."*); Merton wrote that (1936, p. 903) "[…]the "realization" of values may lead to their renunciation", and finally coined the term in his essay in 1948 with the title "The Self-fulfilling Prophecy", referring to Venn and his "suicidal prophecy" as the antonym.

e.g., Ashenfelter and Graddy 2003; 2006; Mei and Moses 2005; Beggs and Graddy 2009), and forecasts based on query reports from Internet searches, such as the Google Flu Trends (Ettredge et al. 2005; Ginsberg et al. 2009; Butler 2008; 2013).

Third, we outline how reactivity could be analyzed empirically by a comparative approach among different markets, which has, to the best of our knowledge, not yet been applied. Finally, we identify how institutions understood as the "rules of the game" can be used to minimize the negative and enhance the positive effects of reactivity (North 1990, p. 3).

In the following section we discuss the literature and background of reactive behavior. Section 5.3 treats the externalities in reactive markets. Section 5.4 provides insights from a comparison of reactive and non-reactive markets. Section 5.5 discusses institutions as a way of coping with reactivity. The last section concludes.

5.2. Background and Literature

As the idea of reactivity to forecasts touches on a large number of topics, we focus this literature review on how the topic has been approached in the social sciences and in particular in economics.[85] As early as 1866, John Venn explained the problem of reactivity: "But when the inference is about the conduct of human beings it is often forgotten that in the inference itself, if published, we may have produced an unsuspected source of disturbance" (Venn 1866, p. 345).

Years later, Morgenstern (1928) adapted the problem of reactivity to economic forecasting and stated his impossibility theorem of a correct prediction in an economic context, where individuals react to predictions. Merton (1936; 1948) coined the idea of a self-fulfilling prophecy in a sociological context, also referring to economic case studies such as bank-runs. Interestingly, he emphasized a reactivity that has a positive impact on the forecasters' accuracy if the market participants were to believe in the forecast strongly enough. This stands in contrast to Morgenstern's claims. Grunberg and Modigliani (1954) showed that Morgenstern's impossibility theorem in a theoretical model (cobweb cycles) does not hold, yet their assumptions have been criticized as partially unrealistic, even by themselves (see, e.g., Henshel 1995). At the same time, Machlup (1955) discussed the problems due to

[85] As Merton (1936, p. 894) stated: "The diversity of context and variety of terms by which this problem has been known, however, have tended to obscure the definite continuity in its consideration."

reactivity from the forecasters' point of view and stated that a verification problem hinders the learning of forecasters. Whenever a model is tested with new data, it remains unclear if the deviations stem from the faultiness of the model or from other unpredictable changes in the system modeled, making systematic model corrections impossible.

Simon (1954) and Baumol (1957) argued that polls and other pre-election predictions can lead to both self-fulfilling and self-defeating prophecies. Both scholars stated that the assumption in their models, which are based on Gruenberg and Modigliani's work, can be criticized in assuming a reaction function of market participants (i.e. in the case of voters) that has to be both known to the forecaster and known to be continuous. Kemp (1962) showed theoretically that correct public forecasting in economics (and other social applications) is impossible. However, his assumptions, in turn, have been criticized by Grunberg and Modigliani (1963). Rotschild (1964) showed that forecasts, although imperfect, can have positive effects in dampening the fluctuations and improving the stability of economic systems.

According to Henshel (1993; 1995) and Hands (1990), the emergence of the idea of (fully) rational expectations (Muth 1961) almost ousted the idea of reactivity. Nevertheless, even under the assumption of fully rational expectations, the quest for reactivity went on. Expectations can be seen as forecasts, thus it is interesting to understand the market outcome to which fully rational "forecasts" would lead. Samuelson explored the problem in 1965, stating that properly anticipated prices fluctuate randomly (i.e. follow a Martingale process). His "random walk hypothesis" became influential especially in relation to financial markets (Samuelson 1965, p. 48; see also Malkiel 1973; Samuelson 1974). The hypothesis states that the reactivity between expectations (i.e. forecasts) and actions in a market leads to a randomly oscillating market outcome. Grossman (1977), among others, showed that reactivity in markets produces "noisy" rational expectations, leading to informational externalities. Grossman and Stiglitz (1980) considered the impossibility of informationally efficient markets, showing that markets therefore can "rest in an equilibrium degree of disequilibrium" (Grossman and Stiglitz 1980, p. 393). The foundation of this idea is close to Rothschild (1964).

Other important issues concerning reactivity and forecasting were raised in the fifties in sociology and social psychology, and in the seventies in economics. Most prominent in economics is Lucas's critique (1976), stating that it is naïve to try to forecast the outcome of a change in policy with a macroeconomic model based on aggregate data. The change in policy would also lead to a change in the microfoundation of the forecasting model, resulting in wrong policy advice. Campbell (1957) and Goodhart (1975) had previously made a similar point, approaching reactivity from the market participants' point of view, who are seen to produce the forecasted outcome: "A reactive measure is one which modifies the phenomenon under study, which changes the very thing that one is trying to measure" (Campbell 1957, p. 298).[86] The manipulation of targets is due to the reaction of the individuals being measured. Thus, the forecaster has to deal with a reactive outcome when trying to predict a new target.

The sociological literature analyzing the problem of reactivity is growing. Since Campbell (1957), it has become a tradition in sociology and social psychology (and psychology in general) to address the problem of the observer and measurement effect as a reactivity problem.[87] Callon (1998) and later MacKenzie (2006) and Esposito (2013) addressed the problem under the term "performativity," which leads to the unpredictability of anthropogenic market outcomes. MacKenzie and Esposito apply their research directly to financial markets. They try to show how the problem of performativity, especially for the field of risk management, has become unmanageable during the recent financial crisis. The very nature of risk management is "that risk can be predicted with probabilistic models" (Esposito 2013, p. 121), but this has been cast into doubt when considering the last financial crisis (MacKenzie 2006, p. 201). In particular, the risk models led to a form of reactive behavior on the part of the market participants (traders) trying to exploit the models (Beunza and Stark 2012). Hence, the models were wrong from the very moment when at least some parameters became known in the market. In this paper, we call such a mechanism "reactivity".

The problem of reactivity involved in forecasts for anthropogenic markets has attracted less attention in the recent literature on economic forecasting (see, e.g., Elliott and Timmermann 2008; Garcia-Ferrer 2012).[88] Certainly, reactivity is always captured in the structural parameters of the forecast models which are discussed in the literature. However, neither the reactivity to forecasts nor possible ways to proxy for the degree of reactivity are dealt with

[86] The idea is identical to the "observer effect" in physics (and indirectly related to Heisenberg's (1927) "uncertainty principle"), which states that by measuring a certain object it is inevitable that the object itself will be affected and, hence, the result will deviate.

[87] In sociology, reactivity is often termed reflexivity or categorized as a form of reflexivity (see, e.g., Espeland and Sauder 2007, p. 6). In particular, researchers must pay attention not to evoke reactive/reflexive behavior in subjects, and thus bias the results of their study (Campbell and Stanley 1963). In financial theory, George Soros (1987) introduced the term "reflexivity" and demonstrated later that he was capable of using it successfully against the Bank of England in the 1992 UK pound sterling currency crisis.

[88] In their review of the literature on economic forecasting, Elliott and Timmermann (2008) do not touch on the problem of reactivity of forecasts; similarly, Garcia-Ferrer (2012) discusses Clive Grangers views on the next steps for a higher predictability in financial market and does not deal with the problem of reactivity neither.

(see also Clements and Hendry 2011).[89] Instead, it is argued that real-time feedback will help the forecasters to learn more quickly and adapt their models faster and more appropriately (Elliott and Timmermann 2008, p. 4). In contrast, we argue that this is unlikely to hold true, because the market participants and their access to new technology (just as the forecasters' access) will enable them to react as fast as the forecasters can recalibrate their predictions.

Another strand of the literature explicitly dealing with a special kind of forecaster, namely security analysts, has so far also neglected the problem of reactivity. Security analysts try to forecast a firm's earnings per share (EPS) over a certain period. Because these forecasts shape expectations on the security market, the managers have an incentive to meet the earnings that have been forecast by the analysts. This leads to a problem of reactivity between the forecast and the actual outcome. From an accounting perspective, the problem of earnings management in firms and public institutions has been analyzed since the early eighties and has produced an abundance of studies (Raman 1981; Healy and Wahlen 1999). The accounting literature since the turn of the millennium has indicated the problem of a possible interaction between analysts' EPS forecasts and the company managers' incentives to converge their earnings towards the analysts' forecasts (see, e.g., Abarbanell and Lehavy 2003; Burgstahler and Eames 2006). Still, Beyer (2008), to the best of our knowledge, is the first study clearly documenting and theoretically analyzing reactivity between analysts' forecasts and managers' earnings management. Yet, the newer literature on analyst forecasts only partly considers the effects of reactivity when calculating analysts' forecast accuracy (e.g., Liu and Natarajan 2012; Bissessur and Veenman 2013).

Similarly, the literature on art auctions generally does not consider the effects of reactivity when treating experts' forecasts about the value of an auctioned artwork (Ekelund et al. 1998 being a notable exemption; for an overview of art auctions, see Ashenfelter and Graddy 2003; 2006). Individual studies have recently opened the field for a more thorough analysis of auctioneer presale price estimates (i.e., art experts' forecasts). These surveys tackle the problem of studying forecast accuracy when the forecast itself can influence the actual outcome (Mei and Moses 2005; McAndrew et al. 2012).

---

[89] For example, we are thinking of ways to proxy the forecasters' reputations and to infer from this the reactivity of the market participants in accordance with the forecast issued. There exist, of course, many studies on learning and the generation of expectation in economic systems (see, e.g., Evans and Honkapohja 2001, for an overview concerning macroeconomics and Wenzelburger 2006, for an overview on mathematical economics), however, they do not study the institutional impact of the group of professional private or governmental forecaster. We focus in this chapter on the impact of these specific groups on market participants and on themselves.

When focusing on new trends in forecasting, we see a similar pattern. A new forecast technique is introduced and the enthusiasm engendered leads people to forget the inevitable reactivity of the market participants. A prominent example is the new promising forecasting technique that uses the flu-related Internet searches to predict flu pandemics (Eysenbach 2006; Ginsberg et al. 2009). However, the problem of reactivity has to be kept in mind. The Google Flu trend forecast during the last winter season (2012/13) substantially overestimated the flu situation in the U.S. (Butler 2013). As argued in Butler (2013), the Google Flu forecast might have suffered from Internet searches looking for news about the flu situation. In addition, the flu hit early and extraordinarily severely last season and, hence, seems to have drawn the public's attention to it. This might have increased the number of searches for preventive measures on the Internet from people who were not ill. Such changes in people's search behavior might have led to a large forecast error for Google's Flu forecast. We assign this effect to the reactive behavior of individuals. But there might be another way in which reactivity can distort such a forecast. If the forecast predicts a severe flu, people might tend to mismatch symptoms of a cold to flu. This would lead to a biased result even if the individuals did not change their searching behavior but, instead, just started Internet searches based on their wrong perception of having flu. Google Flu would be powerless to adapt to this direct form of reactivity, although the company intends to adapt the search-based model to the changes in peoples' search behavior.

5.3. Positive and Negative Externalities of Reactivity

5.3.1. Externalities of Reactivity

From an economic and social point of view, it is important to analyze possible positive and negative externalities of reactivity. Therefore, we first identify the possible courses that reactivity can take and the different outcomes it can lead to:

1. The forecast was ex-ante wrong, but the market participants reacted in such a way that the predicted value occurs. This leads ex-post to the judgment of a correct forecast: Type A (the typical case of a self-fulfilling prophecy).

2. The forecast was ex-ante correct, but the market participants reacted in such a way that the predicted outcome did not occur. This leads ex-post to the judgment of a wrong forecast: Type B (the typical case of a self-defeating prophecy).

3. The forecast was ex-ante correct, and the market participants reacted in such a way that the predicted value occurs. This leads ex-post to the judgment of a correct forecast: Type A.

4. The forecast was ex-ante wrong, and the market participants reacted in such a way that the predicted outcome did not occur. This leads ex-post to the judgment of a wrong forecast: Type B.

In cases 1 and 2, reactivity exerts negative externalities on forecasters and market participants. We focus initially on the market participants who are misguided and who make wrong resource allocations, detecting their error only when the reactive market outcome is corrected, or even never. An example for the first case is a bank run, where a wrong forecast that a bank is not solvent can actually lead to the insolvency of a bank because the forecast causes people to withdraw their money from the bank. Every reactive behavior in accordance with the wrong forecast exerts a negative externality on the other market participants. Merton (1948, p. 194) provides anecdotal evidence that refers to the bank run on the Last National Bank in 1932 as an example of a self-fulfilling prophecy.

It is important to recall that, for example, during a market bubble that emerges due to a wrong, exuberant forecast, the self-fulfilling prophecy can have *positive* externalities in the *short run*. Temin and Voth (2004) provide an excellent example. During the South Sea Bubble, even fully rational investors, informed about the existence of a bubble in the market and with full knowledge about the pessimistic forecasts, still promoted the price exuberance. As long as such rational investors assume that market sentiment is still increasing, they support the further inflation of the bubble, encouraging a self-fulfilling prophecy for a certain period. During this period, the self-fulfilling prophecy exerts a positive externality on the forecaster, who issues the wrong prediction, and the market participants, who believe in this wrong forecast. However, in the long run, when the bubble bursts, the reactive behavior of market participants is unveiled. The negative externalities and the resulting inefficient resource allocation during the bubble become visible.

An example for case 2, a self-defeating prophecy, is a warning about the malfunctioning of a market that is perceived by the market participants as a scare tactic and, thus, strengthens the participants' belief in their misguided actions. Reinhart and Rogoff provide ample anecdotal evidence (2009, p. 208-215). They demonstrate in a compelling way how the "doomsayers" before the financial crisis, e.g., Roubini and Setser (2004) or later Krugman (2007), warned against the vicious circle of borrowing money in steadily increasing amounts in the US. Instead of destabilizing the malign trend, the advocates of "this time it's different" gained ground and the vicious circle continued.[90] Eventually, the predicted outcome did occur. This demonstrates that it is important to focus on a certain period when assigning the various cases to a situation. In the long run, the "doomsaysers" were mostly correct in this episode, while in the short run, they provide a nice example for case 2, of a self-defeating prophecy. In cases 3 and 4, market participants generally face the positive side of reactivity, guiding them to the correct action, or, at least, dissuading them from going in the wrong direction.

Next, we focus on the forecasters' point of view. In case 1, the forecasters that initially issued a wrong forecast but benefit from the reactive outcome that meets their forecast are exerting negative externalities on the forecasters that issued initially correct forecasts. Case 2 mirrors the outcome of case 1 from the point of view of the ex-ante correct forecasters who face the negative externalities of reactivity. Cases 3 and 4 involve no negative externalities. On first sight, the forecasters face only positive externalities, fostering correct forecasts and revealing the wrong ones.

However, a problem still remains with cases 3 and 4. As indicated above, case 3 is ex-post indistinguishable from case 1 (type A problem), and case 4 cannot be differentiated from case 2 (type B problem). This leads to negative externalities imposed by the reaction of market participants. First, for the market participants, both type A and type B problems make it impossible to select the "truly" able forecasters, as ex-ante wrong forecasts can be judged ex-post as correct and vice versa. Therefore, the market participants exert a negative externality on their own future behavior and prevent themselves from learning. Second, the same is true for the forecasters. They will never reliably know if their own forecast was initially correct

[90] Even in past times it was not easy to be a doomsayer. "The [forecasting] activity, induced by the sheer insatiable demand for such forecasts, silenced the pessimists and doomsayers […]", as Bombach (1962, p. 29) explained. (Translated into English by the authors, but stated originally in German: *"Die durch die schier unersättliche Nachfrage nach solchen Prognosen ausgelöste [Vorhersage-] Aktivität hat die Pessimisten und Warner […] verstummen lassen".)*

(type A problem) or was initially wrong (type B problem). Hence, the learning of forecasters in reactive markets is also substantially reduced.

There is evidence that economic forecasts did indeed not improve for different periods and countries, supporting our analysis. Zarnowitz (1992), for example, shows that forecasts for the U.S. GDP did not improve between 1953 and 1989. Heilenmann and Stekler (2012) also show an unsteady behavior in the accuracy of forecasts for German GDP growth between 1967 and 2001 and no trend towards improvement. Despite the massive increase of available data and continuously improving information technologies, there is no discernible trend towards an improvement in economic forecasts (see, e.g., Oeller and Barot 2000; Vogel 2007).

We are not aware of any studies directly addressing the empirical question of the development of forecast accuracy of art experts or security analysts over time. In the case of forecasting techniques using Internet queries, anecdotal evidence exists on the failure of Google Flu trends during the 2012/2013 winter's flu season (Butler 2013). Before that season (2004 to 2008), however, the forecast accuracy of the new Google service did improve the forecast accuracy visibly (see, e.g., Ginsberg et al. 2009). Choi and Varian (2012) provide evidence for other domains, such as forecasts for automobile sales, unemployment claims or choice of travel destinations using web searches from Google Trends. They show that, from 2005 to 2010, most of the algorithms for the forecasts exhibited an improved accuracy. Generally, it is difficult to infer any trends for these new techniques, as they were launched after the turn of the millennium and the short period of time restricts the amount of data available for a reliable empirical analysis on the forecast accuracy over time.

5.3.2. Reactivity-Induced Herding Behavior

The impossibility of learning for forecasters due to self-fulfilling and self-defeating prophecies induced by reactivity leads to several consequences. Firstly, above and beyond the problem of learning, the indistinguishability between a correct and a wrong forecast in reactive markets distorts the forecasters' incentives. It can be credibly assumed that market participants will follow the type A forecasters (the ones that are ex-post correct) when deciding on which forecasters to follow in the future, irrespective of the problem of not being able to distinguish between cases 1 and 3. Hence, the forecasters' reputations rise whenever their forecasts are ex-post judged to be correct. In a reactive market, the forecaster with the highest reputation can induce a self-fulfilling prophecy, so forecasters have a strong incentive to adapt their forecasts towards the prediction of the forecaster with the highest reputation in the market. It is not the true ability of the forecaster (as case 1 and 3 are indistinguishable) but the market participants' reactivity to reputation and the possibility of a resulting self-fulfilling prophecy that create a herding behavior among forecasters.[91]

Secondly, the reactivity-induced herding of forecasters raises two new forms of externalities: The first of these is that each forecaster that starts to herd with the allegedly most reputable forecaster (group of forecasters) in the market increases the pressure on the other forecasters to also follow the increasingly large herd. The second is that the larger the number of forecasters joining the herd, the higher is the false sense of certainty among the market participants and the stronger their reactivity, leading to a self-fulfilling prophecy. Reactivity-induced herding is, thus, a self-reinforcing process.

Several interesting empirical findings support our proposition of reactivity and reactivity-induced reputational herding behavior. For example, an often-cited strand of the literature on forecasting is the literature of security analysts who forecast the earnings per share (EPS) of listed companies (Elliott and Timmermann 2008). Security analysts are identified as herding more when they are younger and when they work with less reputable brokerage houses (Hong et al. 2000; Clement and Tse 2005). Further, Clement and Tse (2005) show a higher accuracy on the part of more experienced forecasters. This result can be interpreted by a reactivity-based explanation, among other explanations, such as reputational herding induced by ability (see footnote 8).

More reputable brokerage houses entice the most reputable forecasters away. So, on average, the analysts with higher reputations work for more reputable brokerage houses and are also older and more experienced, due to the selection process they have undergone (Hong and Kubik 2003). In addition, analysts who are employed by more reputable houses also show a higher relative accuracy, but the accuracy does not increase further once they are employed at the top, i.e., work with the most prestigious brokerage houses (Hong and Kubik 2003; Groysberg et al. 2011). Taken together, the high reactivity in the market for EPS forecasts provides an explanation for such a pattern: Older analysts herd less because they *lead* the herd

[91] It is important to separate our argument of reputation-based herding induced by reactivity from the reputation-based herding induced by ability, as in Scharfstein and Stein (1990) among others. Our argument goes above and beyond the problem that incapable forecasters could mimic the able forecasters. Rather, we state that reactivity can change the outcome such that a wrong prediction by an incapable but ex-ante reputable forecaster may become reality and, thus, will ex-post be perceived as a correct prediction. Likewise, our argument here is not based on a simple informational cascade (Banerjee 1992; Bikhchandani et al. 1992), as the herding forecasters might even know that they are doing wrong, but because the actual outcome is strongly influenced by reactivity (self-fulfilling prophecy), they still engage in herding behavior. All our subjects act in a fully rational manner.

due to their higher reputation. Hence, the reaction to their reputable forecasts induces a self-fulfilling prophecy. Furthermore, once the analysts with a high reputation also work for a prestigious brokerage house and become leading forecasters, they cease to be relatively more accurate than the average forecaster, because the herd approximates their forecasts closely.

The market for EPS forecasts is indeed highly reactive, as the executives have an incentive to manage their companies' earnings in order to meet (or beat) the consensus forecast (the average forecast of the herd). Reactivity-induced herding behavior matches the common empirical evidence in the literature on security analysts. Interestingly, Beyer (2008) neither includes nor touches on the idea of reactivity-induced herding behavior in her model of reactive behavior between the forecasters and the producers of the outcome (managers of the firms), The latter only comprises the reactivity between a single forecaster and a single manager of a firm.

The literature on economic forecasting also provides empirical evidence of herding (Ashiya and Doi 2001; Bewley and Fiebig 2002). However, the object of study in economic forecasts (e.g., predicted GDP growth) is substantially less reactive to the estimates issued than the EPS forecasts issued by security analysts. This can be assumed to apply in the case of large and stable economies. In smaller and more volatile economies, the reactivity to economic forecasts is likely to be higher. Interestingly, this is what the empirical evidence of a cross-country comparison reflects. Forecasts for emerging markets tend to exhibit more herding behavior than forecasts for established industrial or post-industrial economies (see, e.g., Pierdzioch et al. 2012).

5.4. Comparison of Markets With and Without Reactivity

To improve our knowledge about reactivity, we propose to compare forecast markets for reactive and non-reactive systems.[92] A prominent example of forecasts in a non-reactive system is that of weather forecasts. The weather does not react to the forecasts being issued. Self-fulfilling and self-defeating prophecies are not possible in such systems. None of the problems stemming from the indistinguishability between a self-fulfilling or self-defeating prophecy and a genuinely correct or incorrect forecast (type A and type B problems of

[92] Merton (1936) and Morgenstern (1928) mention the difference of forecasts in reactive and non-reactive markets.

reactivity) apply. Hence, we should be able to identify the consequences of reactivity when analyzing the differences between forecasters' behavior in reactive and non-reactive systems.

First, the absence of reactivity is expected to lead to a continuous improvement of forecast accuracy due to the learning of forecasters. The forecasters can (up to a specific measurement error) fully rely on the ex-post measured outcomes, adjust their models, and calibrate their calculations accordingly. The learning of forecasters will establish reliable practices and rules on how to produce correct forecasts.

Second, any technological progress should lead to an improvement of forecast accuracy in non-reactive systems. More data about the initial conditions and the dynamics of the weather as well as advances in the transmission, processing and storage of data will increase weather forecast accuracy over time. There is ample evidence that weather forecasts do steadily improve over time and that learning and technological progress are the drivers of this development (see, e.g., American Meteorological Society 2008; Buizza et al. 2010; Katz and Lazo 2011).

Third, in a non-reactive system, reactive effects such as self-fulfilling or self-defeating prophecies cannot induce pressure on forecasters to herd with the most reputable forecasters. The forecasters would only herd to share the blame; if everybody else fails, the reputational costs fall on all forecasters and remain *relatively* unimportant (Scharfstein and Stein 1990). However, the possibility to distinguish ex-post between a wrong and a correct forecast leads to a correct attribution of reputation by market participants (weather-forecast consumers). Thus, the blame-sharing type of reputational herding behavior will also be driven out in the long run in a non-reactive system. Compared to a reactive system, forecasts for a non-reactive market do not induce a self-reinforcing herding behavior and do ultimately provide a positive incentive to issue the ex-ante most correct forecast possible. We are not aware of any study providing evidence about a possible herding behavior in the market for weather forecasts and, hence, provide an empirical analysis in the following chapter of the thesis testing the propositions stated here.

When we compare the production of a forecast for future weather conditions with that for future economic conditions, both seem to be of similar complexity. As Arrow stated in an interview in 2000 about his military service during World War II as a weather forecaster: “What I found then was another example of a very complex, interacting system. It (meteorology) had a big advantage over economics because the fundamental theory was very well understood."[93] Clearly, the steadily improving weather forecast accuracy attracts the attention of economic forecasters. In a working paper of the Bank for International Settlement (BIS), Vahey and Wakerly (2013) praise the probability forecasting stemming from the meteorological sciences and propose to copy the technique to attain a similar forecast accuracy for macroeconomic forecasting. However, we want to stress the important difference between forecasts in reactive markets and forecasts in non-reactive markets. We perceive the difference between these two types of markets to be neglected in the economic forecasting literature on probabilistic forecasting (Katz and Lazo 2011). The important point is that the market participants in anthropogenic markets not only react to a specific forecast, but also adapt to new forecasting techniques and even adopt them. The reactivity will absorb the improvement of forecast accuracy in economic, financial, or art price forecasting, among many other anthropogenic markets. Any putative increase in forecast accuracy due to more and better data will be neutralized by market participants adjusting their actions with the help of the same new techniques the forecasters are using. Weather forecasts are issued in a non-reactive system and their accuracy and success will, thus, remain distinctively different from forecasting in reactive systems.

Finally, consider the following example. Imagine the case of an unemployment forecast for the next year. The market participants (employees) are assumed to correctly infer from the forecast that they might lose their jobs and, hence, start to work harder and do not ask for higher wages. This reaction results in a self-defeating prophecy for the forecast of higher unemployment. The reactivity might even lead to an increasing pressure to perform among employees.[94] The outcome, though, is desired, and the menace of unemployment is averted. Yet, ex-post the forecasters' prediction is judged negatively, providing a negative incentive for future forecasts of the same type although they are socially and economically desirable. To contrast with the unemployment example, we draw on the story of Joseph in the Old Testament (Book of Genesis). Joseph, after having been asked to interpret a dream of the Egyptian Pharaoh, forecast seven years of abundance and seven years of famine. The Pharaoh had Joseph adopt the necessary measures. Joseph had large storehouses built and filled them with the excess grain during the seven abundant years. Finally, during the seven years of famine triggered by a severe drought, Joseph was able to provide the Egyptians with the grain stored to prevent famine. His own forecast proved to be wrong but was most useful.

---

[93] Quoted according to Katz and Lazo (2011, p. 561).

[94] Lazear et al. (2013) provide evidence that during recessions the productivity of employees increases due to the "work-harder" effect and not due to other effects such as a better selection of workers.

Two issues are important when comparing the examples. First, as discussed above, the forecast for the reactive market of unemployment was self-defeating due to reactivity. An ex-post assessment of the forecast would likely lead to a negative report for the forecasters. In contrast, Joseph was proven right, as the non-reactive weather, not paying attention to the forecast, produced the drought. Second, Joseph used the weather forecast to make another forecast, namely that there would be famine. This second forecast induced Joseph's reaction recommending the storage of grain and, hence, the forecast, by inducing reactivity, ended as a self-defeating prophecy. This example raises an interesting case in which the forecasts for a non-reactive system enter a stage, which we call the anthropogenic level, leading to a reactive system. Forecasts for non-reactive markets are naturally used to forecast probable damage. By doing so, the forecasts are translated into a reactive system. On the anthropogenic level, the forecasts stemming from a non-reactive system face the same problems as forecasts for a reactive market.

Silver (2012) shows that even weather forecasters intentionally bias their weather models to influence consumers. For a large U.S. weather channel, Silver shows that the forecasts issued predicted an excessive probability of rain, but only in a relative way: When the probability of rain was predicted by the initial model to be very low (below 25%), the weather channel issued an *increased* probability. When the initial model, instead, predicted a probability of more than 25%, the forecast was issued without any adaptation. This makes sense, as customers will blame the weather channel for a day when there is no rain predicted and they suddenly find themselves unprotected against sudden rain. Thus, the weather channels impose a "wet bias" and forecast an increased probability of rain although their initial model would objectively predict no rain (Silver 2011, p. 135). The opposite case, when the initial model predicts a high probability of having rain (above 75%), is obviously of less importance for the end-users; they will be prepared for the rain anyway. In contrast to the U.S. weather channels, the U.S. National Weather Service releases almost perfectly calibrated predictions that primarily serve business clients and other weather experts. These examples illustrate that, even for intrinsically non-reactive forecasts, human reactions can affect forecast behavior and accuracy. However, the bias is less severe in such markets than in reactive systems, as the forecast users receive more unbiased information about the true state of the world (see, e.g., Gentzkow and Shapiro 2006).

5.5. Reactions to Reactivity: Institutions

5.5.1. Institutions and Reactivity

Institutions define the "rules of the game" that shape the incentives of economic actors (North 1990; Bowles 1998). Institutions are important for economic prosperity and individual well-being (Frey and Stutzer 2002a; 2002b; Besley 2003; Acemoglu et al. 2005; Frey 2008). Hence, institutions should be used to absorb the negative externalities of forecasts due to reactivity.

Poll restrictions provide an interesting example of an institution that aims at preventing market participants (i.e., voters) from falling prey to reactive behavior.[95] Political scientists discovered that polls before and during elections and ballots may induce reactive behavior on the part of voters, leading to a self-fulfilling prophecy ("bandwagon effect") or a self-defeating prophecy ("underdog effect") (see, e.g., Sudman 1986; McAllistar and Studlar 1991, p. 720). However, the degree of influence of reactivity on voting behavior is difficult to estimate, and thus the implications for public policy are hotly debated (Morwitz and Pluzinski 1996; Feasby 1997; McDonald and Thornburg 2012). Many countries have decided to restrict polls before and during elections and ballots, supporting the notion that forecasts such as polls generate negative externalities due to reactivity. As it is difficult, if not impossible, to internalize these externalities, poll restriction seems a valid instrument for eliminating or reducing reactive behavior. Chung (2012) provides new evidence that 46% of 83 countries[96] surveyed do indeed prevent the release of polls before the elections. The so-called blackout periods range from 1 up to 45 days. Fifty-two percent of the countries report a restriction of polls during the elections or ballots (exit polls).

Another interesting example of an institution that exploits the reactive behavior of market participants is the ranking of colleges and universities in the market for education. During the last three decades, these rankings have become increasingly important for universities in the U.S. and, with a slight delay, in many other countries. Espeland and Sauder (2007) provide an insightful study about the ways rankings induce reactivity. When rankings are perceived as providing useful information about the quality of universities, the best students will be

[95] We interpret polls as forecasts since, from the voters' point of view, polls provide them with information about the most likely outcome of future elections or ballots.

[96] The survey includes the largest and most developed countries on the different continents of the world (Africa (4), Asia (27), Europe (35), North America (8), South America (8) and Oceania (3)). For further details, see Chung (2012, p. 5).

attracted by those that are ranked highest. The highest ranked university can then choose among the most talented students and will hence perform best in the future, as their alumni will perform best in academia and elsewhere.

Frank (1999) argues that the market for education provides an ideal example of a "winner-take-all market". Small shifts in rankings lead to an enormous shift in students' applications. This is reflected in an inefficient proportion of university resources being invested in the factors that are most heavily weighted in the ranking. Hence, even if the forecasters (the producers of the rankings) try to adjust their model to forecast the best universities with an even higher probability, universities will react immediately and adapt to the new rankings (Espeland and Sauder 2007).

The market for education is distinctively different from other product markets, since customers themselves produce the product. Hence, the universities have a high incentive to attract the most able students. This stands in contrast to a sports car manufacturer, to whose reputation the customers' driving skills are of no importance (Frank 1999). The students have a high incentive to see their university at the top of the ranking once they have been admitted. Taken together, the universities, students and forecast (or ranking) providers all share an aligned incentive to perform in accordance with the ranking. This creates a self-fulfilling prophecy.[97]

Why are these examples important? First, they provide insights into how institutions can shape or limit reactivity so as to prevent negative externalities of forecasts. Second, they show how institutions can exploit reactivity in their favor. Third, they make clear that reactivity can drive out competition in forecast markets, giving way to herding behavior. The larger the share of voters who believe in the same pre-election poll, the stronger is the influence of that poll. A stronger poll, hence, leads to a self-fulfilling prophecy by provoking a reactivity-induced herding behavior of the other poll providers towards the first poll, so that they too provide an ex-post accurate poll. A similar process is applicable to university rankings. Interestingly, only the first example has so far provoked an institutional reaction. The reaction leading to a prohibition of polls is reasonable, since the polls do not produce any additional information that is important for the voters' decision making. In contrast, university rankings provide such additional information to prospective students. Rankings fulfill an informational

[97] This is particularly true if the first ranking were so well conducted that the educational landscape was mirrored and the initial ranking matched the perceived reputation among the universities. The highly reputed universities would consolidate their positions from the very first moment of the existence of the ranking.

function in the market for higher education. However, they also pose a problem due to their ability to induce reactivity.

Fourth, the examples emphasize a structural problem of forecasts. Reactivity can give rise to good and bad incentives, and reactivity can produce positive as well as negative externalities. A forecast such as an informative university ranking can prevent market participants (i.e., prospective students), from making wrong decisions, like applying to the worst university and thus wasting time and money. This is an outcome of forecasting in which reactivity leads to a desirable self-fulfilling prophecy. If the forecast were correct and the worst university were to close due to a lack of students applying, the efficient allocation of resources would be increased. However, it is highly doubtful whether an institution can be established that can ex-ante discriminate between the correct forecast (desirable reactivity) and the wrong forecast (undesirable reactivity).

Fifth, the example of the unemployment forecast showed that a self-defeating prophecy can lead to a desirable outcome and prevent market participants from allocating resources to the wrong endeavor. Yet, the problem from the forecaster's perspective is that such useful forecasts can be perceived ex-post as wrong. The forecasters have no incentive to issue such an invidious forecast again. The establishment of a sound institutional reaction to the wrong incentive is difficult. The question is how to restore positive incentives for forecasters that have predicted a bad outcome, which is hopefully defeated and, thus, has led to a self-defeating prophecy, so that the forecast is judged ex-post as wrong.

5.5.2. Potentially Helpful Institutions

Which institutions can be proposed to mitigate the problems of forecasts inducing reactivity? First, we distinguish between forecasts that provide additional information to the market participants (the decision makers) and forecasts that do not. The polls exemplify the circumstances in which there is no provision of intrinsic information, in this case to the voters, while the university rankings do provide information, to a prospective student. It makes sense to restrict or even prohibit forecasts that do not provide additional information can be restricted or even prohibited. Such an institutionalized ban on forecasts may interfere with the political right to freedom of speech. The latter has to be compared with the severity of the forecast's negative externalities that can be prevented with a ban.[98]

[98] For an example, see the discussion of the Canadian poll restrictions in Feasby (1997).

For the second group of forecasts inducing reactivity problems, the informational benefit for the market participants provides no reason to restrict or ban them. For these forecasts, it goes too far to radically conclude, as Morgenstern (1928, p. 121) does, that forecasts have to be banned due to reactive behavior and that only “facts and figures” may be publicized. The problem with issuing *only* raw data is that for most issues an individual cannot make sense of the data without knowing some basics on how to read and interpret them. So, the public always asks for an interpretation and always receives an interpretation, which might be true or not. Furthermore, the increasing information load due to technological progress will not make it easier to keep track of the correct interpretation of specific data. Thus, we have to think about who should provide explanations and interpretations of data, because someone will provide an explanation anyway.

Instead of prohibiting reactivity-inducing forecasts that can provide intrinsic information to the market participants, we propose the re-establishment of an old institution, the devil's advocate. The Catholic Church introduced this institution in the 15th century to prevent clerical decision-makers from choosing the wrong persons to canonize (see, e.g., Stanley 1981). The devil's advocate had the duty of scrutinizing and challenging the reasons advanced for canonizing a certain person. This idea is also an institutionalized form of Janis’s idea that project leaders have to assign the role of “critical evaluators” to every member of a project team (Janis 1972, p. 209).

We propose that the devil's advocate should at least scrutinize the forecast of the most reputable forecaster. The idea is to establish an informational counterweight to the most important factor inducing reactivity in a market. Of course, many questions have to be clarified before the proposed mechanism can be put into practice. In particular, the funding of the devil's advocate and the available resources have to be defined. The most efficient way would be to establish a tax on the forecasters' profit. Such funding would increase with the share and the volume of the specific forecast market, accounting for the degree of the possible external effects. On the other hand, one could argue that the socially desirable effects of such a service (the provision of information as a public good) would justify the use of public funds. A prominent example of such a forecasting institution involves the German Council of Economic Experts, which is nearly independent of the government but is still financed by public funds.[99] The devil's advocate would inform the public about possible arguments against

[99] The German Council of Economic Experts is not completely independent of the German government, as its members are appointed by the Federal President, similar to the U.S. Council of Economic Advisors. However,

the most prominent forecasts and provide information about the data backing the claims. Such a service would result in a less stable environment for reactive behavior as it would interfere with the development of self-reinforcing tendencies. The distortive effect of the devil's advocate on a self-fulfilling or a self-defeating prophecy could be perceived at first sight as a destabilizing effect in a market. However, we expect that such oscillating opinions will be more precise on average in the long run, causing fewer negative external effects than periodically stable equilibria induced by a self-fulfilling or self-defeating prophecy. Whenever the reactivity is interrupted, the putative stable equilibria will experience a fatal market breakdown, leading to a substantive deterioration of markets (see Rothschild 1964, p. 303).

To circumvent the problem of forecasters that ex-ante correctly predict a disastrous future situation but face ex-post negative consequences due to the possibility of a self-defeating prophecy, we propose that an independent forecast provider be established. This proposition can easily be linked to the idea of introducing a devil's advocate forecaster in the market. If the forecasting of negative events does not provide enough incentives, due to the effect of reactivity (self-defeating prophecy), private forecasters will refrain from providing such forecasts. The devil's advocate already bears the duty of publicly challenging the most reputable of the private forecasters, which leads to a provision of information about such imminent negative events. However, we do not believe that such an institution of a devil's advocate would provide correct forecasts. In contrast, we assume them often to be wrong. The important point is to provide a counterweight to the imminence of a self-fulfilling prophecy with harmful negative externalities. The devil's advocate will have a dampening effect on reactivity. The market participants should always be informed about the deficiency of the devil's advocate, not being able to predict the future more reliably than any other forecaster.

5.5.3. International Organizations and Other Official Forecasters as Devil’s Advocates

There already exist arguably independent institutions providing forecasts, such as the International Monetary Fund (IMF), the World Bank, the Bank for International Settlements (BIS), various central banks, and expert advisory councils. A major problem with the

the latter council is distinctly less independent, as it reports directly to the President and issues a yearly “Economic Report of the President”, which is not the case for the German council, which issues a report on its own behalf.

forecasts of these bodies is, however, that their public forecasts have a direct influence on the decision bodies inside the organization itself. The institutions have a substantial capacity to influence the run of a certain economy and thereby induce a self-fulfilling or a self-defeating prophecy by their own actions. Hence, the incentives to forecast against the public prediction of such an influential institution are reduced.

One of the most obvious examples is the Chinese government's forecast of the country's GDP growth and the officially measured growth rates. The incentives for the Chinese government are similar to those of the manager in the case of the EPS estimates of security analysts. It is always best for the managers of the companies covered to meet or beat the consensus expectation of the security analysts. Ever since the Asian financial crises in 1998, the figures officially released on China's GDP growth rate have always overshot the target issued (*The Economist* 2013). But there are also other examples of economic forecasts that are issued by bodies capable of influencing a whole economy: The *World Economic Outlook* of the IMF or the *World Development Report* of the World Bank. Of course, these institutions have to produce forecasts for internal use. However, the publication of their forecasts can lead to a self-fulfilling prophecy, both because of their own actions and also because they may prevent others from issuing deviating forecasts. This risks inducing herding among forecasters and finally among market participants.[100] In particular, smaller and emerging economies that depend more on the forecasts are expected to react more strongly to these forecasts.

A literature has recently emerged that sheds light on these problems. Sockin and Xiong (2013) build on the model of Morris and Shin (2002) on the social value of public information. According to standard economic theory, an increase in the futures price of a commodity should lead to reduced real demand (e.g., Hamilton 2009). Stylized facts from commodities markets show instead that an increase in the futures price may lead to an *increase* in the real demand of a commodity. Sockin and Xiong (2013) provide a model to explain such an effect. Morgenstern (1928, p. 94-95) pointed to this paradoxical effect in his early treatise on forecasts in reactive markets. Sockin and Xiong (2013) are in line with Morgenstern's argumentation. Market participants use the futures price of globally important

[100] Menkhoff and Taylor (2007) and Sarno and Taylor (2001) point to a reactive mechanism by which technical analysis (i.e. the use of e.g., visual inspection of time-series plots or simple moving averages) could lead to reasonable and useful information for predicting exchange rates, although it lacks any economic foundation. They show that, because market participants made use of technical analysis, the central banks also included the technique in their forecast of future movements in the exchange market. By basing their official interventions in the market on such forecasts, the reactive behavior of the central banks induced a precarious situation where official intervention could elevate an irrational method of analysis to a rational one.

commodities to predict the future demand of their customers, resulting in the forecast of their own future demand for the commodities. They start to buy the commodities when (or although) the futures prices are rising. Hence, speculative actions on the futures market trigger inventory market reactions, leading to a reactive effect between the forecast (i.e. the futures price) and the actual outcome. Now, the problem is exacerbated by the fact that economic reports from large organizations, such as the IMF or the BIS among others, have also started to employ commodity prices for public predictions on the future condition of the global economy. This leads to an exacerbated reactive behavior triggering self-fulfilling prophecies and herding behavior among market participants. Ultimately, the herding results in a boom-and-bust cycle of commodity prices, which have actually been experienced in recent years. We conclude that the public forecasts of powerful international organizations should be treated more carefully. Their forecasts provide a substantial source for reactive behavior that potentially leads to distorted markets.

In contrast, the U.S. Federal Reserve System (FED) distinguishes between the release of facts and figures and the release of forecasts in accordance with what Morgenstern (1928, p. 122) recommends for governmental forecasts. The FED only releases its Beige Books (facts and figures) on the current economic condition in a timely manner, but keeps documents containing forecasts (Green Books and Blue Books) for the U.S. and international economy confidential and releases these with a five-year delay.[101] Hence, we suggest that other official bodies should also discriminate between facts and forecasts. The market knows that these organizations will act upon their released forecasts and react accordingly. Forecasts can still be released, but by an independent body, not one incorporating a forecasting unit and a unit that will act directly on the forecasts.

5.6. Conclusion

This chapter sheds light on the problems that publicly issued forecasts can create in reactive markets. Reactivity exists in every anthropogenic market. The issue has been known for a long time (see, e.g., Venn 1866; Morgenstern 1928; Merton 1936; 1948). In these days, however, the stark increase in the amount of available data and in the capacity to generate, transmit, store and process data conjures up the assumption that forecasting in reactive market

[101] See, e.g., the Evaluation Report of the FED's Office of Inspector General from January 18, 2013; Report 2013-AA-A-001.

systems will become more accurate over time. We argue that this is unlikely to materialize because market participants will react to the forecasts and to the forecasters' new techniques. Advances in forecast accuracy will thus be absorbed. We support our claim with empirical evidence from the literature on several areas in economics and the social sciences in general.

We begin by assessing the positive and negative externalities of reactivity, which is induced by any forecast on anthropogenic markets. Further, we present a comparative analysis between reactive market systems and a non-reactive system. We discuss why forecast accuracy regarding non-reactive systems profits from the new achievements in information technology, in contrast to the reactive market system. We draw on evidence in the literature on forecast accuracy of non-reactive weather forecasts to back our arguments.

However, we do not suggest that the use of meteorological models will achieve a higher forecast accuracy in reactive market systems. In contrast, we claim that forecasting for non-reactive markets is distinctively different from forecasting for reactive markets. In a BIS working paper, Vahey and Wakerly (2013) propose that the use of probability forecasting, which is employed in meteorological sciences, would substantially increase the accuracy of economic forecasting. They expect an improvement similar to the accuracy of weather forecasts. In our paper, however, we show why this analogy falls short. The reactive behavior of market participants in the economy and the non-reactive behavior of the weather exhibit a distinct difference. When thinking about the use of forecasting tools and methods from the natural sciences in areas of the social sciences, the difference in reactivity should not be neglected.

Finally, we retrace the ways in which institutions hamper, foster or even exploit the reactivity induced by forecasts. College rankings, as an unusual example of forecasts, represent an institution that changes the rules of the game in higher education by exploiting the reactivity in the market. As another example, we use polls as a type of forecast for voters. Poll restrictions constitute an institution that has been set up as a reaction to the reactivity of voters, because their reactive behavior can lead to a self-fulfilling prophecy or self-defeating prophecy in elections.

Merton (1936, p. 894) noted in his seminal article on reactive social behavior that the "diversity of context and variety of terms by which this problem has been known, however, have tended to obscure the definite continuity in its consideration." It is surprising how the idea of reactivity still appears in various individual areas of the social sciences *without* the similar roots of the problem being remarked upon. Thus, a comparative analysis is important for gaining further insights into how forecast users react to the forecast and how this reaction influences the outcome predicted.

As a final example, not even the pursuit of happiness is free of problems due to reactivity. Schooler et al. (2003) argue that the individual pursuit of happiness can be self-defeating from a psychological perspective. To forecast individually that performing a particular act will make us happy often turns out to be a fallacy. On the aggregate level, Frey and Stutzer (2002a; 2009), and Frey and Gallus (2012), argue, drawing on empirical evidence, that the maximization of national happiness can have negative effects for the population. For example, when it becomes an official target to increase a national happiness index, the reactive behavior of politicians, bureaucrats and citizens could lead to biased incentives, manipulations and misrepresentation in a national happiness index. We, hence, expect that the pursuit of a national happiness forecast would induce strong reactive behavior with similar effects to those of electoral polls. They may be both self-fulfilling and self-defeating. While the use of happiness research is important when considering institutions that can help individuals to become happy, the pure ranking and seeking of national happiness might have detrimental effects due to reactivity. Even in the most human quests, we should not neglect reactivity when publicly releasing a forecast.

6. Forecasts and Reactivity: A Comparative Analysis of Three Markets With Different Levels of Reactivity

In this chapter, we provide empirical evidence for the theoretical reasoning on forecasts and reactivity in the preceding chapter. Reactivity distorts the forecasting of anthropogenic market outcomes, as the market participants record and process a new forecast and react according to its information value. By doing so, these market participants naturally influence the outcome that the forecast is intended to predict. To investigate empirically how reactivity affects forecasting, we use large data sets to compare forecasts in three areas with different levels of reactivity, namely the areas of meteorology, finance, and the arts. We empirically analyze three main elements of the theoretical reasoning of the preceding chapter. Firstly, we intend to show that forecast accuracy increases over time for non-reactive outcomes, but does not do so for forecasts on reactive outcomes.[102] Secondly, we provide evidence about how reactivity is influenced by fundamentals and market sentiment and about how temporal changes of reactivity can affect ex-post measured forecast accuracy in reactive markets. Thirdly, we aim to demonstrate empirically how the different levels of reactivity in markets can induce different behavior on the part of forecasters, leading to a herding behavior among forecasters that vitiates forecast accuracy.

6.1. Reactivity and Changes in Forecast Error Over Time

In reactive markets, such as the stock or the art market, forecasters cannot use measured outcomes as unbiased reference points to calibrate their models. Thus, forecasters in reactive markets face the problem of a target that is not only moving exogenously but that is also distorted endogenously by their own earlier input.

We hypothesize that in reactive markets neither the advances in information technology and data processing nor the learning of forecasters have a positive long-term effect on the

[102] For an extensive explanation of the concept of reactivity, see the introduction to the preceding chapter.

predictability of outcomes. First, all market participants in these anthropogenic markets adapt to the new technology or the new data processing techniques. Second, the same is true for learning, as in a reactive market it is not only the forecasters who learn, but also the market participants. This impedes forecasters' learning and its positive effect on forecast accuracy.

In the case of financial markets, security analysts forecast company earnings of stocks they cover and are employed at brokerage houses or brokerage units that can be part of a larger bank or a financial services company. Today, analysts can rely on a much larger amount of macroeconomic and firm-level data and on a much faster information flow than a decade ago.[103] On the other hand, due to the faster information flow, managers of companies that are covered by the analysts also learn much more quickly about the issued forecasts and, thus, can react to them much faster. Likewise, managers learn about new data sources and the new knowledge the analysts base their forecasts on and adapt to such advances as well. This results in a zero sum game. Bai et al. (2012) provide supporting evidence for this hypothesis. They find that, although the efficiency of information production in financial markets has improved over the last fifty years, the usefulness of stock and bond prices in forecasting earnings of companies has been stable over time.

We see a similar mechanism in the art market. The fast and easy access to Internet databases of the historical art market prices of various providers (e.g., Artnet.com, Artprice.com, Artfact.com among others) has improved the informational environment drastically (Horowitz 2011, p. 211-12). This improvement, which started in the early nineties, had a massive impact on the informational resources of individual art collectors and less experienced professional investors as they gained better and faster access to information about the past market prices of art objects. Before, only the old and established auction houses had access to most of these data, as collecting information about auction prices and estimates was always one of their important tools against other market participants. Most of the art market data was collected by the auction houses themselves; the larger and more established the house, the more auctions it held, hence the more precise the data received from its own sales. Old auction houses such as Sotheby's and Christie's still have an advantage over the data on old paintings that are

[103] In a conversation with two older security analysts, they stated that during the nineties, they still cut out newspaper articles by hand and collected them in folders of the companies they covered. They used the hard copy editions not because they did not have access to computers but because the newspaper articles were not readily available online.

exceptionally rarely sold.[104] The Internet databases often cannot provide any auction data from earlier than the seventies.

However, if the new information on the Internet can reduce the information asymmetry between forecasters and bidders, the latter will become more independent of the auction house's expert knowledge. The reactivity between bidders and the art experts' forecasts will be reduced as a result of the better informational environment for bidders. As a result, today's art market prices should exhibit more anchoring on historical sales prices than before the drastic increase of information via the Internet. Beggs and Graddy (2009) provide evidence that anchoring on historical art prices existed between 1980 and 1994, before the rise of Internet art data bases. They show that anchoring took place on both sides, among buyers and in the auctioneers' presale estimates. Yet, the authors are not able to establish whether the auctioneers merely reacted to the buyers' anchoring, or whether they were affected by anchoring themselves.[105]

Forecasting of future outcomes in the reactive art and financial market is in stark contrast to forecasting in a system in which no reactivity exists, namely the weather.[106] The weather outcome is not influenced by the forecast, and the forecasters can make best use of this lack of reactivity. Forecasters can genuinely calibrate their models against the unbiased outcomes. Thus, we expect forecasters competing in the market for weather forecasts to be able to steadily improve their forecast in two different ways: Firstly, by continuous learning that improves their knowledge of the interactions and interdependencies of meteorological forces, and secondly, by advances in information technology (timely transmission, more data storage capabilities, faster data processing systems) and in the generation of big data (better sensor technology, more observation stations, weather satellites, ships and aircrafts). Both of these will have a direct, positive and steady impact on weather forecast accuracy. Thus, we predict that the innovations in information technology and the increase in available data (quantitative and qualitative) will continuously improve the forecasts in the weather market.

[104] We were told by a managing director of one of the two most famous auction houses that these houses still have an advantage due to their extensive price archives when it comes to the presale valuation of an exceptionally rarely traded old artwork.

[105] "Thus, while the auctioneers' and sellers' behavior may exhibit reference point effects in choosing the low estimate and secret reserve, it does so in a way consistent with buyers' behavior" (Beggs and Graddy 2009, p. 1037).

[106] Please also refer to Section 5.5 of the preceding chapter.

An intuitive approach to measuring the improvement of forecasts is to calculate their forecast error (FE), which equals the difference between the forecast and the actual outcome (forecast – actual = FE). Taken together, we state the following proposition:

Proposition P1: Despite the positive technological development and learning of forecasters in all three markets, only the error of weather forecasts permanently and constantly decreases over time, while the forecast error in financial and art markets does not continuously improve.

6.2. How Can Reactivity Be Identified in a Forecasting Market?

To identify reactivity between forecasts and the actual outcome in a market is by definition a very challenging task due to its endogenous characteristic (see Chapter 5). Ideally, we would have data about the identical market situation occurring twice, once with and once without a forecast being issued. The difference between the two market outcomes would be due to the reactivity of the market participants induced by the forecast.

As we are not aware of any data available from such an extraordinary situation, other methods to identify reactivity have to be found. First, one could try to reconstruct the outcome of a situation with a forecast as if there no forecast had been issued (i.e., to reveal the outcome without reactivity). Second, situations could be found in which, due to exogenous shocks, the level of reactivity in a market is altered. Such situations would provide an opportunity to explore the impact of different levels of reactivity on the forecast error. As we aim to compare the different markets, we do not pursue the first method. It seems impossible to find a single or closely comparable technique to reconstruct the "true" counterfactual in the art and financial markets. Furthermore, the current literature on financial accounting is rather arbitrary in the construction of the counterfactual, i.e., the reconstruction of "true" and unmanaged earnings (see, e.g., Cohen et al. 2008; Beyer et al. 2010; Simpson, forthcoming). The literature on the economics of art is mostly silent about the issue.[107] Instead, we focus on situations which are assumed to alter the level of reactivity in a market, thus resulting in a predictable and directly measurable change in forecast errors. To do so, we first have to identify factors that could exogenously change the level of reactivity in a given market.

[107] Mei and Moses (2005) provide a notable exception.

Therefore, we scrutinize the incentives of the forecasters and the market participants who originate the actual market outcome.

In the case of earnings per share (EPS) forecasts issued by security analysts in the financial market, the forecasters' career concerns provide incentives to produce low forecast errors (see, e.g., Hong et al. 2000, and the discussion of the analyst literature in Chapter 4 of this thesis). On the other hand, managers of stock companies that are covered by analysts share a high incentive to meet (or slightly beat) the expectation of the market. This expectation is usually represented by the average opinion (consensus) of the EPS forecasts issued by the analysts. The EPS consensus for a specific firm represents one of the most important accounting figures by which the executives' and the firms' performance is measured (see, e.g., Matsumoto 2002; Beyer 2008). If this threshold is not reached, the valuation of the firm, the costs of capital, and the executives' compensation through stock or option plans are negatively affected. Hence, the managers have substantial incentives to meet or beat the expectations of the market (i.e. the analysts' consensus) and will react and adjust the firm's EPS wherever possible. Several studies document the effect of earnings management of firms that try to reach a certain level of EPS, mostly by over- or understating accruals in their books (see, e.g., DeGeorge et al. 1999; Cheong and Thomas 2011). An increasing strand of the financial accounting literature supports the reasoning that managers in general aim to issue "smooth earnings" over the years; such behavior is perceived as a safe and reassuring sign by investors (see, e.g., Kirschenheiter and Melumad 2002, p. 762; Repenning and Henderson 2010, p. 1). Therefore we can expect the managers to have the largest incentives to meet or just marginally beat (by one or two cents) the consensus forecast (see, e.g., Simpson, forthcoming). Combining these incentives yields a strong reactivity between the forecasts and the market outcomes in the case of EPS estimates.

In the case of the auctioneers' presale price estimates for art works, forecasters' career concerns provide incentives to produce low forecast errors, too. The art buyers' incentive is to buy the artwork at the lowest price possible (just marginally higher than the willingness to pay of the second highest bidder). However, the art experts' forecasts can still provide guidance to the art buyer or bidder, either by anchoring or as a reference point (see, e.g., Beggs and Graddy 2009).[108] Mei and Moses (2005) provide evidence that the price estimates

[108] Often, the terms are used interchangeably. Yet, for example, Kahneman (1992) uses *reference dependence* as an effect that impacts the reference point in the assessment of gains and losses when they are valued asymmetrically, *anchoring* as the effect that influences judgment of what is normal more generally.

of auction houses influence the actual prices. They focus on the market segment of expensive paintings, as this segment is the most profitable in terms of commissions for the auction houses. The authors conclude that auction houses might rationally try to exploit credulous investors. This finding is in contrast to the outcome of Milgrom and Weber's (1982) model. They state that in a competitive market setting with rational agents, the auction houses with an "honesty is the best policy" strategy make the most profits and, thus, should issue honest price forecasts. Ashenfelter (1989) and McAndrew et al. (2012) support the model with empirical evidence. On the other hand, Ekelund et al. (1998), Beggs and Graddy (1997; 2009) and Mei and Moses (2005) provide evidence that supports the idea that auction houses can and sometimes do exploit their influence on fetched prices by overestimating their price forecasts. The estimates also constitute a reference point for the reservation price, which is usually around 80% of the auctioneers' lower estimates and can also influence the bid of a potential art buyer at an auction (Ashenfelter and Graddy 2006).[109]

We conclude that it is reasonable to assume that (1) reactivity exists in the art market because auctioneers' presale price estimates influence the art buyers, in particular if the buyers are poorly informed, or credulous, due to anchoring or by providing a reference for them, but that (2) this reactivity is substantially lower than in the case of EPS forecasts in financial markets, as the buyer in principle shares no incentive with the auctioneer to buy the work for a high price; instead, the art buyer intends to buy at the lowest price possible.

After having analyzed the incentives of the forecasters and of the market participants, who finally produce the outcome in the different markets, we are able to identify the factors influencing these incentives. We assume the forecasters' incentives to be stable over time and concentrate our analysis on the outcome producers, i.e. the art buyers and the managers. First, we focus on the financial sector and the EPS forecasts.

6.2.1. Market for Earnings per Share Forecasts (Financial Market)

Drawing on the model by Beyer (2008), we state that the managers' reactive behavior is determined by the cost that the managers have to bear if they do not meet the analysts' consensus. These costs define the level of reactivity in the market for EPS forecasts. The individual manager's costs are influenced by three factors (Beyer 2008):

[109] Auctioneers usually issue a presale price range estimate for an artwork, defined by an upper and a lower bound. If the artwork does not reach a bid higher than 80% of the lower bound (sometimes also 70%, see, e.g., Ashenfelter and Graddy 2011), the work is not sold but "bought in" by the auction house.

1. Economic resources of the managers' firm

2. Market pressure on managers to fulfill expectations

3. The penalty and probability of detection of earnings management and financial fraud due to laws, regulations and investigative authorities.

1. The firm has to have resources sufficient for the manager to manage its earnings up to the threshold of the consensus forecast. These resources depend on the condition of the firm, which in turn depends on the general economic situation. If the general economic situation deteriorates, the firm's resources will decline to the point where the managers, even if fully exploiting earnings management, cannot generate enough earnings to meet the consensus.[110] The general economic situation provides the base for the firm's earnings as well as an *upper bound* for the earnings management: During harsh economic periods the capacity of managers to engage in earnings management is reduced, resulting in a decreased reactivity in the market for EPS forecasts. However, during flourishing economic conditions, an asymmetry in the incentives for the managers prevents them from beating the consensus forecast even though they might have the resources to do so. This asymmetry lies between the managers' costs if not meeting the market's expectations and the benefits if they beat the expectations. Beating the expectations bears strongly decreasing marginal returns. As investors generally reward smooth earnings development over time (Kothari 2001; Kirschenheiter and Meluad 2002), managers only have limited incentives to overshoot the analysts' consensus, even during booming economic periods. Without reactivity in the market, the analyst would face the same difficulties in forecasting a correct EPS during a booming market period and during a busting market period. However, the reactivity in the market for EPS forecasts induces the asymmetry that during flourishing periods the forecast errors are low, because the managers have the highest incentive to meet the forecast although they could release higher earnings. Only during a deteriorating market period, when the firms' reduced resources do not allow managers to meet the analysts' consensus, is the forecast error expected to increase due to the lower reactivity.

[110] This holds unless the managers are willing to engage in blatant financial fraud, and their costs increase dramatically, as they might have to face litigation and personal prosecution (see, e.g., the cases of financial fraud in Dyck et al. 2010).

2. The market pressure on managers is another important factor in inducing or relaxing reactivity in a market. If the investors' perceptions of the market are bullish (i.e., a strong upward trend is expected), managers' costs of not meeting the EPS consensus will be high. During a bearish period, when the investors' sentiment is low, the managers' costs will be lower when not meeting the analysts' consensus (see, e.g., Simpson, forthcoming). Hence, increasing investor sentiment (orthogonalized to the general economic conditions) leads to higher costs for managers if they do not meet the consensus forecast and, finally, leads to higher reactivity.

3. Obviously, the degree of the expected penalty and the probability of detection of the earnings management influence the managers' costs. The probability and the degree of the penalty represent the negative incentive in regard to the managers' costs if not meeting (or beating) the consensus. The lower the expected costs from the penalty, the more the managers engage in earnings management, i.e., the higher is the reactivity in the market.

Considering these three factors, we define the following propositions and recall that reactivity is assumed to lead to lower forecast errors:

Proposition P2: Due to the asymmetric incentives of managers to meet the analysts' consensus, reactivity induces larger forecast errors in the financial market during strong general economic downturns than during market upturns of the same intensity.

Proposition P3: Due to the asymmetric incentives of managers to meet the analysts' consensus, reactivity induces larger forecast errors during times of decreasing investor sentiment than during increasing investor sentiment of the same intensity.

The third factor of the managers' costs if not meeting the consensus forecast also leads to a proposition; however, we can state at this point that we are not able to test Proposition P4. This is for two reasons. First, we are not aware of data or of an index that could be used to track the influence of expected penalties and their probabilities on the executives' costs in managing earnings over the period between 1999 and 2010. Second, the passage of the Sarbanes-Oxley (SOX) Act in 2002 seems to mark an important shock to the costs of earnings management, as it imposed stricter rules for financial accounting in corporations. However, it remains unclear if SOX finally led to a net increase in costs, as empirical evidence merely shows a substitution away from the stronger regulated accrual-based management in financial accounting statements towards the real earnings management by adjusting discretionary spending in R&D, advertising or the other activities of firms (see, e.g., Graham and Harvey

2005; Cohen et al. 2008). Therefore, we assume the costs due to potential detection and punishment of earnings management to remain constant for the managers over the period between 1999 and 2010. For the sake of completeness, we still present *Proposition P4: In addition to the influence of the general economic condition (P2) and the investor sentiment (P3), an increase in the expected penalty and/or the probability of detection of earnings management leads to a higher forecast error as reactivity is reduced.*

6.2.2. Market for Presale Price Forecasts of Artworks (Art Market)

In contrast to the market for financial forecasts, the market for art forecasts exhibits a much lower reactivity, due to the buyers' incentive to buy at the lowest price possible. Buyers generally have no incentive to follow the auctioneers' price forecast. Hence, the reactivity in the art market must be due to behavioral effects such as anchoring (Beggs and Graddy 2009) or credulity due to the limited attention of art buyers (Mei and Moses 2005). The same factors as in the case of the financial market have an impact on the art market:

1. The general economic situation is also relevant for the art market and provides an upper bound for reactivity, as the buyers cannot spend more money on a certain artwork than they possess. However, Renneboog and Spaenjers (2013) and Dimson and Spaenjers (2011) provide evidence about the lower correlation between the art market and other economic indicators than the correlation of the stock market and the same indicators. Similarly, while in the case of the EPS forecasts, most managers have only a low incentive to exceed the analysts' consensus forecast (see, e.g., Kirschenheiter and Melumad 2002), there is no reason why art buyers should perceive the auctioneers' estimate as an upper bound and not outbid the estimate. We conclude that the market for presale art price forecasts has a reduced reactivity compared to the market for financial EPS forecasts.

2. In the case of the art market, buyers experience no direct market pressure to achieve a certain level as is the case of the consensus forecasts for managers in the financial market. Yet, market sentiment does naturally influence the individual buyer, though not in a way that affects reactivity, as it does not influence the buyer's behavior towards the auctioneer's estimate. In the art market, changes in investor sentiment generally increase the difficulty to produce accurate pre sale forecasts. In contrast to the financial market, the difficulty in the art market increases with the *intensity of the change*, no matter in which direction. Thus, substantial increases and declines in investor sentiment are both expected to raise the forecast

error in the art market. (Recall that in the financial market only a negative change in investor sentiment is expected to lead to a higher forecast error due to the decline in reactivity.)

3. As no institutions such as regulations or other accounting rules exist to define how an art work has to be valued and booked, there is no such influence on the art buyers' incentives and, thus, no influence on reactivity.

We conclude that the factors used to analyze the changes in reactivity in the financial market do not have the same impact on the reactivity in the art market. We are not aware of factors that could be used to analyze exogenous shocks to the art buyers' anchoring or the limited attention of art buyers at an aggregate level and not only at the individual level. Hence, at the aggregate level, we assume reactivity to exist in the art market and to be stable and substantially lower than in the financial market.

Therefore, the forecast error behaves distinctively different for art price estimates than for financial EPS forecasts due to the different levels of reactivity and its symmetric reaction to changes in art investor sentiment. In the financial market, forecasters have to bear larger forecast errors when the reactivity collapses due to a negative change in general economic conditions or investor sentiment; a positive change is absorbed by the reactivity of the market because the managers have the incentive to meet or only marginally beat the consensus forecast. This stands in contrast to the art market, in which the art buyers have no such asymmetric incentives. The art buyers do not aim to meet the auctioneers' presale estimates, and they do not face asymmetric costs for under- or overshooting the estimate. Thus, in the art market, the forecast errors increase with a change in art investor sentiment, no matter in which direction, whereas in the financial markets the forecast errors increase only due to negative changes. We condense these thoughts on the art market in our propositions:

Proposition P5: In the weakly reactive art market, the changes in general economic conditions will have a much lower influence on the forecast error than in the financial market.

Proposition P6: In the weakly reactive art market, the stronger the change in art investor sentiment is, the higher is the forecast error, independently of the change's direction.

6.3. Data

To test our propositions, we rely on three data samples from the markets for weather, art and financial (EPS) forecasts. All of them include forecasts and actual measured outcomes to calculate the forecasting quality in the form of a forecast error.

6.3.1. Weather Data

To analyze weather forecasts, we use data from the Swiss National Weather Service (Federal Office of Meteorology and Climatology – MeteoSwiss).[111] Each day at noon, MeteoSwiss issues its main weather forecast for the next day. Therefore, the weather service produces detailed forecasts about the relative sunshine duration, precipitation, wind and temperature extremes in various regions of Switzerland. They retrieve data from their own weather stations spread all over the country and from other weather services that produce regional and global weather data models, such as the European Centre for Medium-Range Weather Forecasts (ECMWF) or the U.S. National Weather Service. To keep our analysis as comprehensible as possible and not to have to rely on partly arbitrary classifications of hit rates for categorical meteorological parameters such as classes of sunshine duration, we only retain forecasts for the maximal temperature of the following day in degree Kelvin in our sample. In Switzerland, as in most other countries, this temperature forecast is economically important for the prediction of future energy consumption. Utility companies base their expected demand for electricity to heat and cool offices, apartments, storage rooms, hospitals, and factories on such estimates.

MeteoSwiss also provided the actual outcomes for each day from the installation of the OPKO System (in German: *Objektive Prognose Kontrolle*; in English: Objective Prediction Control) on January 1, 1999 until the last revision of the system on December 3, 2010. For a reasonable coverage of the climate in Switzerland, we received the data for six different regions in Switzerland, four in the Swiss midlands (Zurich, Basel, Bern, Geneva), one in the southern part of Switzerland (Lugano), and one in the Swiss mountain region (Sion). With this selection of regions, we are able to cover most of the climatic and topological environments for weather forecasting in Switzerland. This leaves us with a sample of 4,352

[111] We are highly indebted to Ludwig Zgraggen, Alexander Giordano, Christophe Voisard and Jaques Ambuehl for making the data available and for their invaluable help in understanding and compiling the data in a form that is usable for our econometric purposes. Furthermore, they provided important insights into the nature of weather forecasting.

observations for each of the six regions, which produces a total of 26,112 observations. Unfortunately, MeteoSwiss was not able to retrieve the forecasts for some time periods due to wrongly coded database entries. The periods concerned are July 2000 to July 2001, October to November 2001, and November to December 2007. We have to exclude all observations with missing forecast data, as we need this information for our analysis. Finally, we arrive at a data sample with 22,454 observations from MeteoSwiss, consisting of a forecast of day t for the temperature maximum on the following day (t+1) as well as the actual measured maximum temperature of the following day (t+1).

6.3.2. Financial Data

For our sample of forecasts and the actual outcomes in the financial market, we rely on the earnings per share (EPS) forecasts of sell-side analysts for the companies of the U.S. stock market (NASDAQ and NYSE). We retrieved the data from the Institutional Brokerage Estimation System (I/B/E/S) Detailed Earnings History Files. The database was accessed through the Wharton Research Data Services (WRDS) interface. This database provides the yearly EPS estimates of almost all analysts covering U.S. stocks and working for a brokerage house in the US. The houses include such large financial services firms as Citigroup or Morgan Stanley which employ more than 100 analysts but also small investment advisers with only one or two analysts. The EPS are forecast for a specific financial year of each company. We only retrieved observations for companies that have their financial year ending at December 31, which is the vast majority of U.S. firms, and we only consider forecasts that were issued less than 12 months before the financial period end. We focused on the same period as with the other data samples, namely on the years between 1999 and 2010. This provided us with 1,078,380 observations. Lastly, we discarded 32,265 (2.9%) observations that either have a missing forecast or a missing actual EPS in the database. This left us with a final data sample of 1,046,115 observations with EPS forecasts and actual EPS, including 763 brokerage houses employing more than 11,036 analysts covering 7,831 individual stocks.

6.3.3. Art Data

For the data on art forecasts and the actual prices fetched at the auctions, we rely on a data set covering the vast majority of Chinese art auction sales between 1994 and 2011 including more than 735,593 individual lots from 7,143 auctions that took place in 426 auction houses mostly in China (including Hong Kong, Macau and Taiwan) but also in the UK and the USA;

the total sales turnover aggregate to more than RMB 88 billion, which equals USD 14.2 billion (at 2013 values of RMB of 6.20 per USD). The data were obtained from www.artron.net, one of the largest databases covering Chinese artworks (Bai et al. 2013).[112]

The data consist of classical Chinese paintings, calligraphy and oil paintings. For our final sample, we only retain lots, which have an auctioneer presale price estimate (the forecast) as well as a hammer price (the actual outcome) for the artwork in the database.[113] Furthermore, to have the same period as for the weather or financial data, we focus on auctions between 1999 and 2010. This reduces our final art data sample to 323,478 individual lots (i.e. observations) that took place in 262 auction houses and aggregate to a total sales turnover of more than RMB 57 billion.

During the last decade the Chinese art market experienced a boom during our sample period and has advanced to one of the economically most important art market places in the world. In Mainland China, the sales revenue rose from RMB 97 *million* in 2000 to RMB 32 *billion* in 2010 (equals more than USD 5 billion at 2013 values of RMB of 6.20 per USD).). *Art Market Trends*, the annual art report 2011 from Artprice, identified Beijing as the global top market place for art in terms of revenue covering more than 27% of the global art auction revenues (see also, Bai et al. 2013; Horowitz 2011).

6.3.4. Descriptive Statistics

Table 6 provides the descriptive statistics of the forecasts, the actual outcomes and the forecast errors of the three forecast markets. Panel A shows the statistics for the weather data sample with 22,454 observations. On average, the daily maximum temperature measured in Switzerland is 288.7 in degree Kelvin (K), which corresponds to 15.6 degree Celsius (°C).[114] The minimum of the daily maximum temperature between 1999 and 2010 in one of the six regions was 264.2K (-9.0°C) and the maximum was 314.2 K (41.0°C). The median maximum temperature was 289.2K (16.0°C).

The forecast error is measured in three different forms: (1) *Temp_FE* is the raw forecast error (FE) and is defined as FE = forecast - actual, i.e. *Temp_FE* = *TempForecast* - *TempOutcome*;

[112] We are thankful and highly indebted to Jia Guo, who originally collected the data, for sharing her dataset with us.

[113] Prices and estimates in other currencies than the Chinese Renminbi are converted with the daily average exchange rate of the auction day accessed through Bloomberg).

[114] Kelvin is converted to Celsius in a simple formula; Kelvin = Celsius + 273.15. This is due to fact that Kelvin defines the absolute zero. In degree Celsius this minimum is measured as –273.15C.

(2) *Temp_AFE* is the absolute raw forecast error and is defined as *abs(Temp_FE*); (3) *Temp_AFEP* is the absolute forecast error in percentage of the actual outcome and is defined as *Temp_AFE*/abs(*TempOutcome*). The forecast errors in panels B (Art data) and C (Financial data) are calculated in the same way.

The meteorologists on average underestimate the maximum temperature but still score a low mean raw forecast error (*Temp_FE*) of -0.3K (or °C) and a median raw forecast error of -0.4K. The outliers reach from a raw forecast error of -14.2K to an overestimation of 41.5K. The absolute raw forecast error (*TEMP_*AFE) has a mean of 1.7K, a median of 1.3K, a natural minimum at 0K and a maximum again of 41.5K. Next, we focus on the forecast error that is most comparable among the different markets, that measured in percentage of the actual outcome (*AFEP*). The *Temp_AFEP* shows a mean error of 0.58% and a median of 0.45%. The minimum naturally lies at 0.00% and the maximum at 15.2% with an interquartile range from 0.21% up to 0.79%. The 99th percentile stands at low 2.21% and the 1st percentile still at 0.00%.

Panel B of Table 6 shows the financial data sample, containing EPS forecasts of the U.S. stock market with more than a million observations. The average EPS forecast of USD -124.25 and a median of only USD 1.20 falls in line with the analysis that the distributions are highly skewed. Both the *EPSForecast* and the *EPSOutcome* show extreme outliers in their distributions, with minima at USD -3,767,500 and -2,275,000, respectively, and maxima at USD 180,000 and 178,000, respectively. However, the interquartile ranges start at USD 0.41 and 0.32 and end at USD 2.35 and 2.30 for the forecasts and the actual outcomes, respectively.

We observe skewed distributions for the forecast errors (*EPS_FE*); the median *EPS_FE* is exactly 0, though its mean is 17.12. The maximum *EPS_FE* is USD 2,117,500 and the minimum USD -1,492,500, while the interquartile range is between USD -0.10 and USD 0.17. The absolute raw forecast error (*EPS_AFE*) has a median of 0.12 and a mean of USD 51.08. Its interquartile range is narrow, starting at USD 0.04 and ending at USD 0.39. The 5th percentile is USD 0.01 and the 95th percentile is USD 2.35. The absolute forecast error in percentage of the actual EPS (*EPS_AFEP*) has a corresponding median of 9.71%, a mean of 55.7% and an interquartile range from 3.1% to 30%. The maximum lies at the extreme of 435,900%. It is important to bear in mind that the skewed distribution of the forecast errors strongly suggests the use of the median whenever the errors are aggregated.

Finally, the art data sample (Panel C, Table 6) includes more than 320,000 observations. On average, the auctioneer presale forecast of the sale price (*PriceForecast*) is RMB 113,564 (18,317 USD)[115] with a minimum of RMB 6 and a maximum of RMB 403 million. The auctioneers usually provide an upper and a lower bound estimate to give the unskilled and inexperienced retail client an estimated price range rather than a precise point estimate. However, in order to arrive at a forecast that can be compared with the outcome, we use the average estimate between the upper and the lower bounds as the point forecast (*PriceForecast*). This procedure is standard in the literature on art economics (see, e.g., Mei and Moses 2005; Ashenfelter and Graddy 2006). The median lies at around RMB 20,000, much lower than the mean, which suggests a highly right-skewed distribution, as expected when it comes to auction prices (estimates) of artworks. The actual prices fetched (*PriceOutcome*) at the auctions show a similar picture. The mean price notes at RMB 176,877, the minimum price is RMB 100, and the maximum is RMB 169 million.

While here the maximum price forecast exceeds the maximum actual price by almost 240%, the interquartile range for the actual prices between RMB 9,900 and RMB 89,600 lie much closer to the range of the forecast prices from RMB 7,000 to 62,500 and surmount the forecasts. The auctioneers on average underestimate the sales price of artworks, which can be seen in the distribution of the raw forecast errors *Price_FE* that have a negative mean (-63,313) and a negative median of -3,440. Further, the large difference between mean and median of the raw forecast error suggest a highly right-skewed distribution. The same skewness appears when the absolute raw forecast error is employed, as the mean *Price_AFE* lies at RMB 73,613 while the median is RMB 6,200. Still, when focusing on the minimum and maximum raw forecast error (*Price_FE*), the auctioneers overestimated the actual sales prices by a large extent (RMB 375 million), whereas they underestimated it by only RMB 144 million. When focusing on the absolute forecast error as a share of the actual outcome, the *Price_AFEP* is right-skewed with a mean of 34.8% and a median of 25.8%. Again, these numbers strongly suggest employing the median as the aggregation instrument of choice when comparing forecast errors in the different markets.

[115] Aggregated figures in Chinese Renminbis (RMB) in this data description section are always converted at 2013 values of RMB of 6.20 per USD unless stated otherwise.

Table 6: Descriptive Statistics of Weather, Financial and Art Data

Panel A: Weather Data (N=22,454)

Variables	Mean	Min	1th	5th	25th	Median	75th	95th	99th	Max
TempForecast	288.4	264.2	271.2	274.2	281.2	289.2	295.2	302.2	306.2	314.2
TempOutcome	288.7	263.1	270.7	274.3	281.7	289.3	295.8	302.4	305.8	311.8
Temp_FE	-0.3	-14.2	-5.6	-3.8	-1.7	-0.4	0.9	3.0	5.1	41.5
Temp_AFE	1.7	0.0	0.0	0.1	0.6	1.3	2.3	4.3	6.3	41.5
Temp_AFEP	0.0058	0.0000	0.0000	0.0004	0.0021	0.0045	0.0079	0.0150	0.0221	0.1522

Panel B: Financial Data (N= 1,044,288)

Variables	Mean	Min	1th	5th	25th	Median	75th	95th	99th	Max
EPSForecast	-124.25	-3767500.00	-9.50	-1.50	0.41	1.20	2.35	6.25	53.40	180000.00
EPSOutcome	-141.37	-2275000.00	-14.16	-2.15	0.32	1.15	2.30	6.14	54.00	178000.00
EPS_FE	17.12	-1492500.00	-3.35	-0.75	-0.10	0.00	0.17	1.60	10.50	2117500.00
EPS_AFE	51.08	0.00	0.00	0.01	0.04	0.12	0.39	2.35	16.00	2117500.00
EPS_AFEP	0.5574	0.0000	0.0000	0.0039	0.0308	0.0971	0.3005	1.7500	7.4625	4359.6230

Panel C: Art Data (N= 323,478)

Variables	Mean	Min	1th	5th	25th	Median	75th	95th	99th	Max
PriceForecast	113,564	6	464	1,500	7,000	20,000	62,500	400,000	1,650,000	403,000,000
PriceOutcome	176,877	100	880	2,300	9,900	30,800	89,600	599,040	2,576,000	169,000,000
Price_FE	-63,313	-144,000,000	-1,060,000	-210,000	-21,800	-3,440	380	10,000	57,515	375,000,000
Price_AFE	73,613	0	0	200	1,600	6,200	25,840	220,049	1,086,390	375,000,000
Price_AFEP	0.3478	0.0000	0.0000	0.0182	0.1071	0.2584	0.5058	0.8140	0.9495	908.0909

Notes: The table reports summary statistics for the weather data sample (Panel A), the financial data sample (Panel B), and the art data sample (Panel C). In all three samples we retain only observations with non-missing forecast and actual outcome. The variable *TempForecast* is the forecast of the maximum temperature during the following day *t+1* issued on day *t* measured in degree Kelvin (K=°C+273.15). *TempOutcome* is the actual measured maximum temperature on day *t+1* for which the forecast was issue in *t*. *Temp_FE* is the raw forecast error defined as *(TempForecast - TempOutcome)*. *Temp_AFE* is the raw absolut forecast error defined as the absolut value of *Temp_FE*. *Temp_AFEP* is the absolut forecast error in percentage of the outcome, defined as (*Temp_AFE*/abs(*TempOutcome*)). *PriceForecast* is the auctioneer presale price estimate of an artwork. *PriceOutcome* is the actual price fetched of the artwork at the respective auction. *Price_FE, Price_AFE* and *Price_AFEP* are defined analogously to *Temp_FE, Temp_AFE* and *Temp_AFEP. EPSForecast* is the earnings per share (EPS) forecasts issued by a sell-side analyst during year *t* targeting the EPS of a specific company *j* at the end of the financial year *t*. *EPSOutcome* is the actual EPS figure made public by company *j* for the financial year *t*. *EPS_FE, EPS_AFE* and *EPS_AFEP* are defined analogously to *Temp_FE, Temp_AFE* and *Temp_AFEP*.

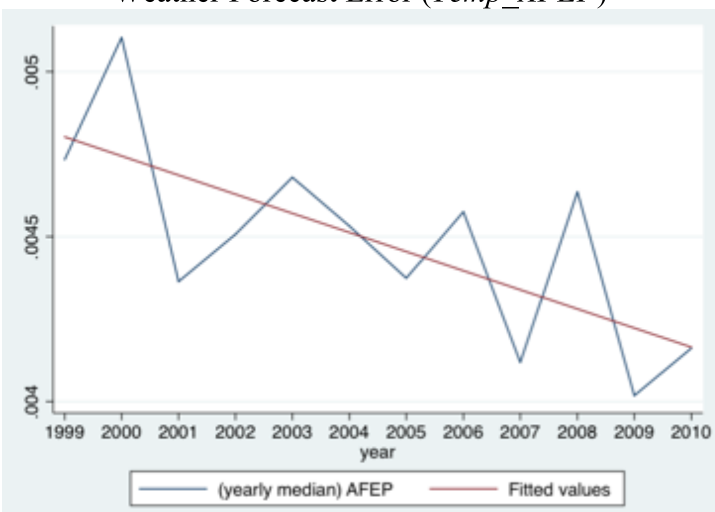
Data Sources: MeteoSwiss, I/B/E/S, www.artron.net

6.4. Empirical Findings

In Proposition P1, we argue that, despite the positive technological development and learning of forecasters in all three markets, only the quality of weather forecasts constantly improves over time while forecasting quality for financial and art markets does not improve (or only temporarily and unsteadily). Figure 5 provides evidence that supports our proposition. The development of the yearly median of the absolute forecast error in percentage of the actual outcome (*AFEP*) steadily decreases in the sample of weather forecasts (*Temp_AFEP*).[116] For the other reactive forecast markets, finance (*EPS_AFEP*) and art (*Price_AFEP*), no improvement of the forecast error is visible over the years from 1999 to 2010. The fitted regression line in each of the figures only has a statistically significant *negative* coefficient (-0.00006; t-value = -3.04) in the case of the weather forecast. This confirms the proposition that the forecast error for weather forecasts decrease over time. The fitted lines for the art and the financial forecasts bear both a positive coefficient, and the coefficient for financial EPS forecasts is even statistically significant at the 10% level.[117]

Figure 5: Yearly Median of the Absolute Forecast Error in Percentage of the Actual Outcome (*AFEP*) for Weather, Art, and Financial Forecasts

Weather Forecast Error (*Temp_AFEP*)



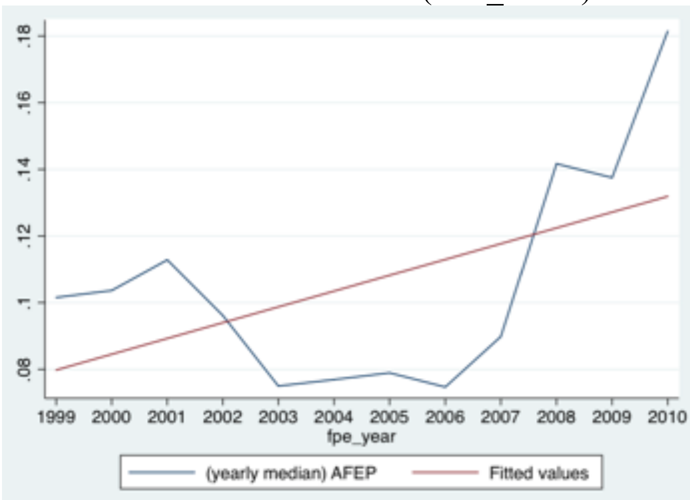
Notes: The fitted line has a negative coefficient of -0.00006, which is statistically significant at the 1% level (t-value= -3.04); for details, see Appendix C Table C1. Data Source: MeteoSwiss

[116] If not specified differently, throughout our analysis we use the median to aggregate the forecast error per year, as we have shown that the forecast errors have highly skewed distributions. However, we obtain the same results if we use the mean to aggregate the forecast error.

[117] In the robustness section of this chapter, we provide additional evidence for a negative stationary trend in the weather forecast time series and a nonexisting trend in the art and financial forecast time series.

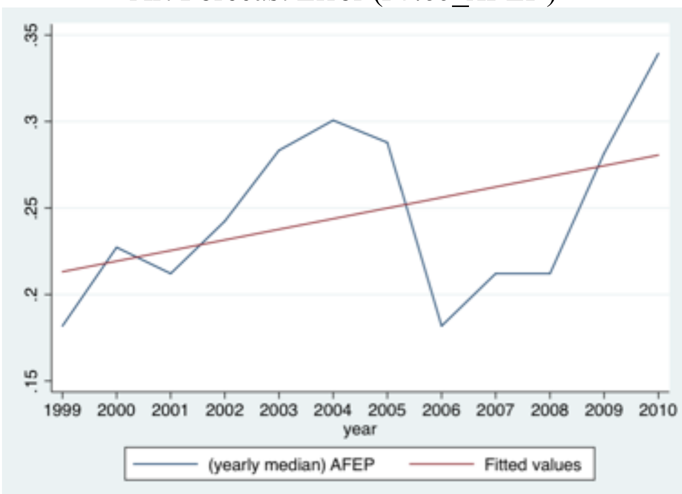
(Continuation of Figure 5)

Financial Forecast Error (*EPS_AFEP*)



Notes: The fitted line has a statistically significant positive coefficient of 0.005 (t-value= 1.92); for details, see Appendix C Table C1. Data Source: I/B/E/S

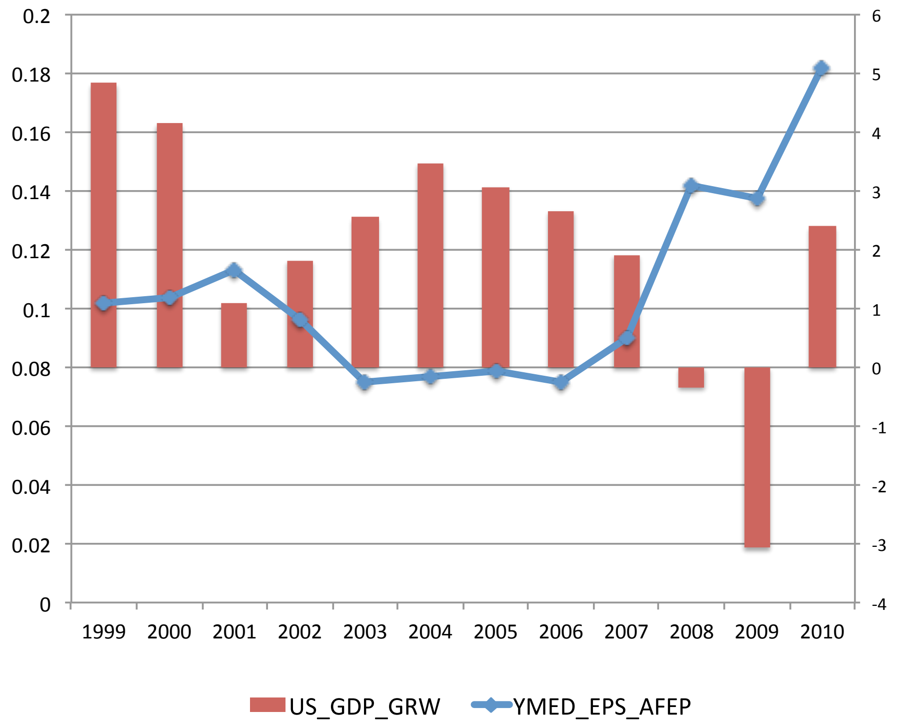
Art Forecast Error (*Price_AFEP*)



Notes: The fitted line has a statistically insignificant positive coefficient of 0.006 (t-value= 1.54) ; for details, see Appendix C Table C1. Data Source: www.artron.net

Our next step is to identify the existence of reactivity in the art and finance forecast markets and to provide evidence for different levels of reactivity in the two markets. We propose in Proposition P2 that, due to high reactivity, large aggregated forecast errors occur in the financial market only during general economic downturns. During economic upturns, the high level of reactivity leads to low forecast errors. In principle, without reactivity, the difficulty in forecasting the EPS during a strong downturn should be as difficult as during an economic upturn of the same intensity. However, Figure 6 provides evidence supporting the idea of asymmetry in forecasts errors between economic up- and downturns. We see that the forecast error (measured as the yearly median *AFEP*) is lowest during periods of substantial positive GDP growth, as in 1999 and 2000, and between 2003 and 2006.[118] Whenever the economy is cooling down or contracting, the forecast error increases. However, the relation does not always fit perfect, especially not for 2009, when the U.S. economy faced an immense contraction but the forecast error was not increasing. Instead, it was slightly decreasing at a high level. However, we have proposed two mechanisms that influence the reactivity in the market for EPS forecasts. After focusing on the general economic condition, we now turn to the second one, investor sentiment.

Figure 6: Financial EPS Forecast Errors and U.S. GDP Growth Rate in Percentage Year-On-Year



Notes: Financial EPS forecast errors (yearly median *EPS_AFEP*, left axis) and U.S. GDP growth rate in percentage year-on-year (right axis). Data Sources: IMF World Economic Outlook database, I/B/E/S.

In Proposition P3, we mention investor sentiment that is orthogonal to the general economic condition as the other important factor determining the reactivity of EPS forecasts (Baker and

[118] We used the data on U.S. GDP growth from the IMF World Economic Outlook database (http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/index.aspx), accessed on July 12, 2013.

Wurgler 2006; 2007).[119] A positive change in investor sentiment is expected to put pressure on the managers to meet the consensus forecast, thus increasing reactivity and leading to a lower forecast error. A negative change, instead, is expected to lead to a higher level of forecast error. Managers face lower pressure to meet the consensus forecast and the resulting lower reactivity leads to a higher forecast error. Figure 7 confirms such a relationship between investor sentiment and the EPS forecast error. Interestingly, the link fits in each year, even in 2009. This strongly supports our notion of high reactivity in this market, as investor sentiment is calculated so as not to have any direct influence on the firms' earnings, unlike the change of GDP (Baker and Wurgler 2007).

Figure 7: Financial EPS Forecast Errors and Yearly Growth Rate of Baker and Wurgler's Investor Sentiment Index

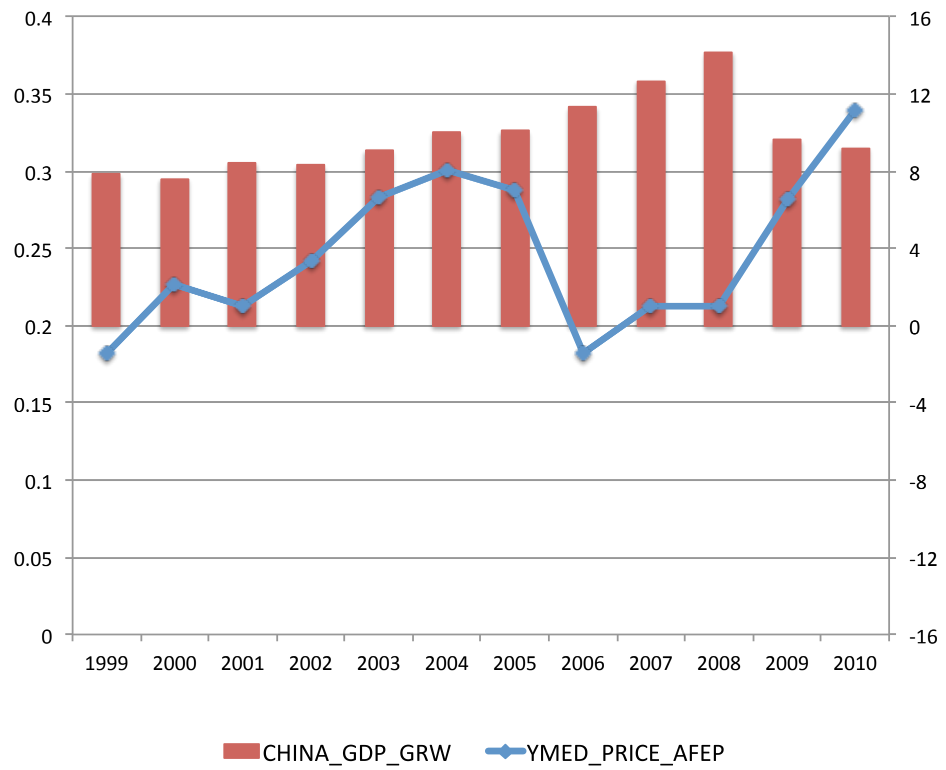
Notes: Financial EPS forecast errors (yearly median *EPS_AFEP*, left axis) and yearly growth rate of Baker and Wurgler's investor sentiment index (*INVSENT*, right axis). Data Sources: Investor sentiment index (Baker and Wurgler 2007; updated 2011 by Jeff Wurgler), I/B/E/S.

---

119 We use the data on investor sentiment from Jeff Wurglers NYU homepage (http://people.stern.nyu.edu/jwurgler/), accessed on July 12, 2013. Specifically, we employ in our figures the yearly average of the monthly investor sentiment data that is orthogonalized to macroeconomic factors, as developed in Baker and Wurgler (2007). Since 2007, the data is updated by Jeff Wurgler until January 2011.

Next, we compare these results from the financial market forecasts with the art market forecasts data. First, in Figure 8, we see no visible link between the auctioneers' forecast errors and the Chinese percentage change in GDP[120] as suggested in Proposition P5, because art prices tend not to react substantially to GDP (and other macroeconomic shocks, such as firms' earnings, see, e.g., Dimson and Spaenjers 2011; Renneboog and Spaenjers 2013).

Figure 8: Art Auctioneers' Forecast Errors and the Yearly Growth Rate of the Chinese GDP



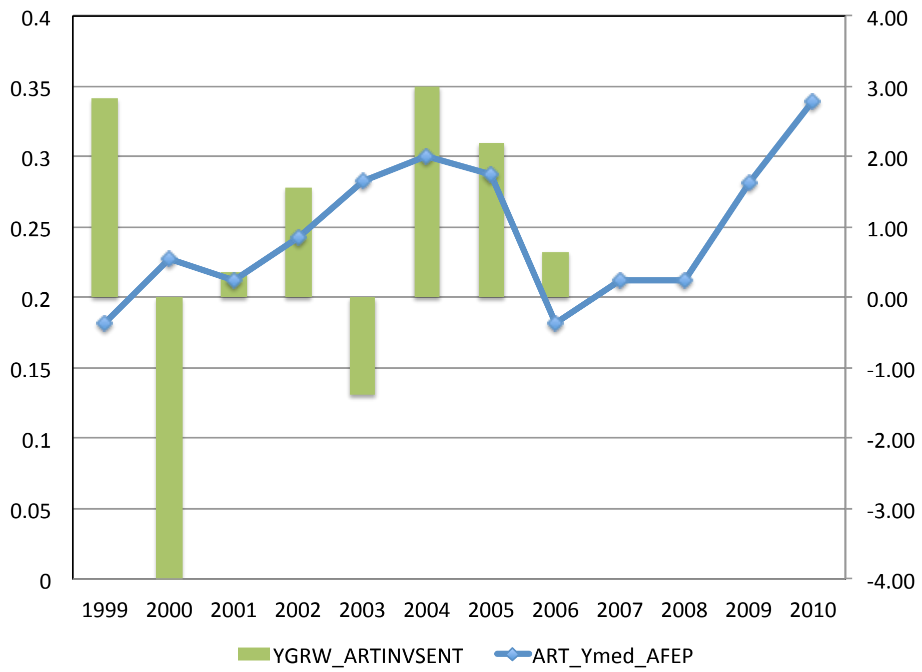
Notes: Art auctioneers' forecast errors (yearly median *PRICE_AFEP*, left axis), and the yearly growth rate of the Chinese GDP (*NGDP_RPCH*, right axis).Data sources: IMF World Economic Outlook, www.artron.net

Again, we add a proxy for art investor sentiment (Proposition P6), as was the case with the financial market EPS forecasts. To the best of our knowledge, there exists no art market sentiment index tailored to the Chinese art market, so we employ the global art market investor sentiment index by Renneboog and Spaenjers (2013).[121] Unfortunately, the index runs only until 2006.[122] Yet, in Figure 9, it becomes visible that there is a link between the

[120] We used the data on China GDP growth from the IMF World Economic Outlook database (http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/index.aspx), accessed on July 12, 2013.
[121] The data on the art investor sentiment index is used as depicted in Renneboog and Spaenjers (2013, p. 51).
[122] Because we do not have access to the underlying data, we could not extend the index over our full period.

forecast error and the global art market sentiment, as proposed in Proposition P6: Only the *intensity* of the changes in sentiment determine the level of the forecast errors and not the direction of the changes in sentiment. The rationale is simple – if sentiment changes and market changes are large, the actual market outcome is more difficult to predict than when sentiment remains stable. This can be seen in 2001 and 2006, when an extraordinarily low change in art investor sentiment leads to a lower level of forecast errors than in the year before.

Figure 9: Art Auctioneers' Forecast Errors and the Yearly Growth Rate of Renneboog and Spaenjers' Art Market Sentiment Index



Notes: Art auctioneers' forecast errors (yearly median *PRICE_AFEP*, left axis) and the yearly growth rate of Renneboog and Spaenjers' art market sentiment index (*YMED_PRICE_AFEP*, right axis, no data available after 2006).Data sources: Art investor sentiment index (Renneboog and Spaenjers 2013), www.artron.net.

Recall that the high level of reactivity of the EPS forecasts leads to lower forecast errors only when the change in the investor sentiment is positive but not when it is negative. This asymmetry does not exist in the art market, because the art buyers have no incentive to smooth the auction prices, resulting in a reduced reactivity in the art market. Hence, we have provided evidence supporting the existence of high reactivity in the market for financial EPS forecasts and a low reactivity in the market for auctioneer presale art price estimates. The degree of reactivity in combination with the surrounding incentive-set in a market has an empirically measurable influence on forecast errors and the development of forecasting quality over time.

6.5. Institutional Factors and Reactivity-Induced Herding Behavior

Institutions and the resulting incentives shape the behavior of forecasters, similar to North's (1990, p. 3) notion of institutions as "the rules of the game". We focus on an important institutional factor in virtually every market, namely on reputation. A forecaster's reputation is certainly among the most important institutional factors (besides technical resources and access to data) affecting the environment and the incentives for forecasters in a specific market. Reputation shapes the way market participants react to forecasts. We assume that market participants react the more strongly to a forecast the higher the reputation of the forecaster in a specific field.

Whenever a forecaster or a specific group of forecasters learns that it bears the highest reputation in the market, it can exploit that reputation and try to induce a self-fulfilling prophecy (SFP) to its advantage by biasing the forecasts. In the art market, auctioneers might use their high reputation to influence the art buyer with an overestimated presale price estimate so as to fetch higher prices and, thus, generate more commission. In the financial market, a brokerage house might use its high reputation to issue an overestimated EPS forecast for a specific stock to support clients that are invested in that stock company. The reputation of the forecaster in combination with the overoptimistic forecast would support the positive view in the market. If the reputation is strong enough and the company has enough resources that its managers can react to the forecast and meet it, the reactivity induces a self-fulfilling prophecy. At the same time, the most reputable forecasters (i.e. the *reputation leaders*) face the problem of avoiding damage to their reputation while they are exploiting it, as it might always happen that a manager (i.e. a company) cannot reach a biased consensus forecast.

First, however, in a highly reactive market, the bulk of less reputable forecasters will also learn about the mechanism through which the forecasts of the reputation leaders have the highest probability of becoming self-fulfilling in a highly reactive market. They will start to adjust their forecasts towards the reputation leaders. Such a *bad learning* of forecasters induces a herding behavior that conceals a possible forecast bias of the reputation leaders and dampens their risk of reputation damage as the herd provides protection from public accusation. In the case where the forecast is ex-post judged to be wrong, the reputation leader can hide in the herd. The majority of the other forecasters (i.e. the herd) will be wrong, too, so the reputation leader bears no reputational damage. In the case where the forecast was biased but induced a self-fulfilling prophecy, so that the forecast of the reputation leaders is ex-post judged to be correct, their reputation will even increase.

Second, in a less reactive market, such a mechanism is less pronounced. The less reputable forecasters learn that, due to the lower reactivity, the reputation leaders are less able to induce a self-fulfilling prophecy. Thus, the less reputable forecasters share a much lower incentive to follow the reputation leaders' forecasts, resulting in less herding behavior and a higher risk for the reputation leaders to bias their forecasts. The herd of less reputable forecasters will not provide the reputation leaders enough protection to hide in the herd, and the absence of the herd impedes the inducement of a self-fulfilling prophecy. We condense these considerations in two propositions:

Proposition P7: In a highly reactive market (e.g., the market for financial EPS forecasts), the less reputable forecasters learn about the most reputable forecasters' capability to induce a self-fulfilling prophecy and herd with the reputation leaders' forecasts.

Proposition P8: In a market with low reactivity (e.g., the market for presale art price estimates), the reputation leaders cannot induce (as often as in the financial market) a self-fulfilling prophecy and, thus, the less reputable forecasters do not (or much less) herd with the leaders.

When we compare the market for financial forecasts and the market for art forecasts, we arrive at the following predictions. In the case of financial EPS forecasts, we expect the higher reactivity to induce a substantial herding behavior among the less reputable forecasters towards the forecasts of the most reputable forecasters. Hence, there will be no visible difference between the forecast error of the group of reputation leaders and the median forecast error of all the forecasters.

In contrast to the financial market, the substantially lower reactivity in the art market induces a much lower herding behavior of the less reputable forecasters towards the reputation leaders. Thus, we expect (1) a visible difference between the forecast errors of the reputation leaders and the forecast errors of the median forecaster, and we expect (2) that the average forecast error of all forecasters is lower than the forecast error of the group of most reputable forecasters. Point (2) is because lower reactivity allows less reputable forecasters to issue a forecast that is less biased towards the reputation leaders' forecasts. All these forecasts that are less biased, although not fully independent, provide a case for the wisdom-of-crowds effect (Galton 1907; Surowiecki 2005). We expect the median forecast error (yearly median *AFEP*) among all forecasters to be lower in the less reactive market for art forecasts. For the financial EPS, we predict no such difference. The median forecast error of all forecasters and the median forecast error of the reputational leaders are expected to be congruent. We anticipate that the herding behavior of less reputable forecasters weakens the wisdom-of crowds-effect by inducing a strong interdependence between forecasts (Lorenz et al. 2011).

It is important to make clear that we are not able to directly measure herding behavior in this setting. In the financial EPS forecast market, we could use the same herding measure as in Chapter 4 of this thesis, where we define herding behavior as the case when a forecaster revises his prior EPS forecast towards the herd (i.e. the EPS consensus of all the other security analysts). In the art market, where usually only the auctioneer publicly issues a forecast for a specific artwork, we are not able to build a consensus forecast for the same artwork without more than one forecast. Hence, we cannot apply the same measure as in the case of EPS forecasts. In contrast, in the art market, some authors create an artificial consensus price forecast for a specific artwork based on a hedonic art price index (Ashenfelter and Graddy 2006). They finally compare this artificial consensus forecast with the auctioneers' forecast to show biased forecasting behavior (see, e.g., Mei and Moses 2005 or Beggs and Graddy 2009). Such a measure is not (readily) applicable to the financial market, as the construction of a comparable hedonic price index and the resulting artificial consensus forecast would be perceived as arbitrary. Hence, we rely here on the temporal changes in the development of forecast errors of different groups. However, we argue that our comparative analysis of the different changes in forecast errors is able to show that herding behavior is the most obvious explanation for the patterns of the forecast errors and that the level of reactivity can explain the degree of herding behavior in a market.

Finally, one could argue that the congruence of the median forecaster and the reputation leaders' forecasts is not due to the herding behavior of the less reputable forecasters in a highly reactive market. Rather, it might be because all the forecasters share exactly the same information, the same information processing techniques based on the same theoretical background. In this case, reputation should play no role, and all the forecasters that have similar resources and access to data should exhibit a similar forecast error. We can test this reasoning if we can identify a group of forecasters that shares an almost identical access to resources and know-how but is of different reputational market power. This third group of forecasters, which we call the *challengers*, would share a high incentive to reach the status of a reputation leader. They could do so by partially deviating from the reputation leaders' forecasts (that also becomes the herd's forecast) trying to achieve a lower forecast error in the cases of companies that could not react and meet the reputation leaders' forecasts.

As we assume substantial herding behavior in the market for EPS forecasts, due to the reactivity in the market inducing the reputation leaders' capability to provoke a self-fulfilling prophecy, we expect the challengers group of forecasters to achieve a lower forecast error than both the reputation leaders' forecasts and the median forecasts. If there were no herding behavior, the lowest forecast error should be achieved by the median forecast. Therefore, in a market with low reactivity such as the art market, we expect only low herding and, thus, we predict that the median forecast error of all the forecasters is lowest due to the wisdom-of-crowds effect. We summarize this in two propositions:

Proposition P9: In a highly reactive market, the herding behavior of the majority of less reputable forecasters increases the median forecast error and facilitates a partially lower forecast error of the challengers group of forecasters that are close to the status of reputation leaders and of similar resources and know-how.

Proposition P10: The same (P9) is not true in markets with low reactivity, where the median forecast error of all forecasters is lower than the forecast error of any other group of forecasters.

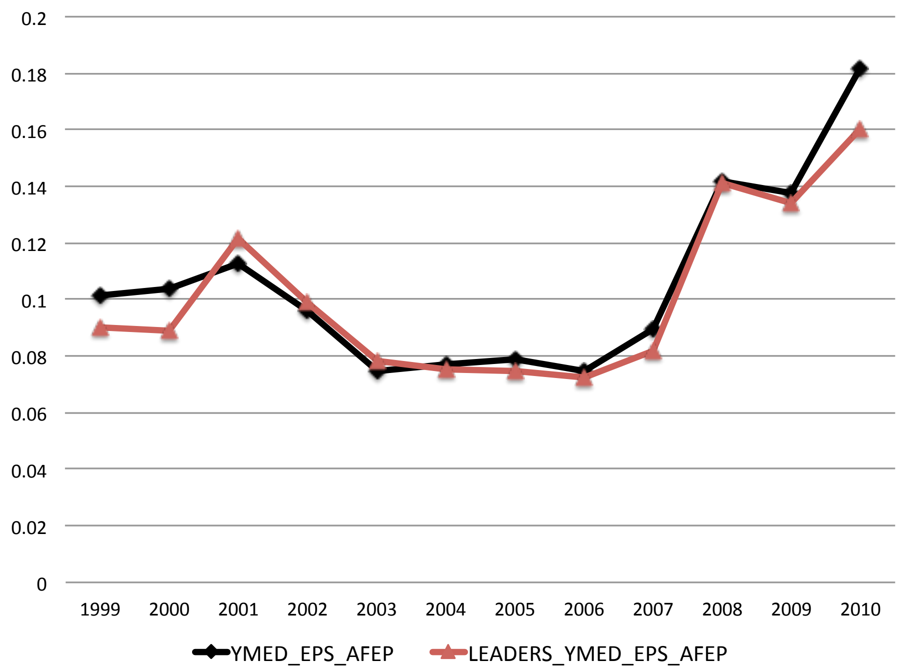
6.5.1. Empirical Evidence

To test these predictions, we first have to identify the most reputable forecasters in the two markets. In the Chinese art market, there seems to be a rather clear ranking order among the auction houses, with four houses leading at the top (Bai et al. 2013; Horowitz 2011, p. 13; Boll 2011; Esman 2012): The domestic Poly International Auction and the China Guardian, and the international art auction giants Christie's and Sotheby's (mostly operating in the Chinese market from their Hong Kong-based subsidiaries). Even on a global scale, in 2011 these auction houses occupied the first four places in the rankings of the renowned survey *Art Market Trends 2011*. Focused on the Chinese art market from 1999 to 2010, China Guardian achieved total sales of RMB 9.579 billion, Poly RMB 6.426 billion, Christie's RMB 5.210 billion and Sotheby's RMB 3.893 billion. For the analysis, we group the forecast errors of these four Chinese art market giants and use their yearly median *AFEP*. We compare their AFEP to the yearly median *AFEP* of all auction houses in the data sample, a total of 262 individual auction houses.

In the financial EPS forecast market, we identify the reputation of a brokerage house by counting the number of analysts a certain house employs. This reputation measure is in accordance with the literature and usually called "broker size" (see, e.g., Hong et al. 2000; Clement 1999; Hilary and Hsu 2013). Then, we rank the brokerage house according to their average broker size between 1999 and 2010 and identify the four largest houses, similar to the art market. These four brokerage houses employ a large number of analysts, from 120 analysts to 142 analysts on average over the 12 years. For the analysis of the forecast error, we group the forecast errors of the four most reputable brokerage houses in the U.S. market and present the yearly median *AFEP*. We compare this forecast error to the yearly median *AFEP* of all brokerage houses in the data sample with a total of 763 individual brokerage houses.

Figure 10 provides evidence in favor of the effect expected in Proposition P7. The yearly median forecast error (*AFEP*) of the group of reputation leaders and the median error of all forecasters in the market show a very congruent course over the years. We regard this as supporting evidence for our proposition that in a highly reactive market the forecasters exhibit a herding behavior towards the forecasts of the reputation leaders. In contrast, Figure 11 shows a starkly diverging trend of the median forecast error (*AFEP*) of the reputation leaders and the yearly median forecast error of all forecasters in the art market, as suggested by Proposition P8. The low reactivity exhibits a marginal pressure on the less reputable forecasters to herd with the reputation leaders. Additionally, due to low herding behavior, we suggested that the more independent forecasts lead to the wisdom-of-crowds effect, which induces a lower forecast error for the larger group of less reputable forecasters. This picture is confirmed in Figure 11; the median forecast error (*AFEP*) of all forecasters is substantially lower than the error of the reputation leaders.

Figure 10: Financial Market – Yearly Median Forecast Error (*EPS_AFEP*) of All Forecasters and of the Reputation Leaders in the Market for EPS Forecasts



Notes: Yearly median forecast error (*EPS_AFEP*) of all forecasters (*YMED_EPS_AFEP*) and of the reputation leaders (*LEADERS_YMED_APS_AFEP*) in the market for EPS forecasts; 1999-2000 Data source: I/B/E/S

Figure 11: Art Market – Yearly Median Forecast Error (*PRICE_AFEP*) of All Forecasters and of the Reputation Leaders in the Market for Art Price Forecasts

Notes: Yearly median forecast error (*PRICE_AFEP*) of all forecasters (*YMED_PRICE_AFEP*) and of the reputation leaders (*LEADERS_YMED_PRICE_AFEP*) in the market for art price forecasts; 1999 – 2010. Data source: www.artron.net

The challengers group of forecasters has to be identified to provide evidence about Propositions P9 and P10. In both markets, the forecasters in the third group should have a marginally lower reputation than the reputation leaders in the market but nonetheless share a similar access to the technical and informational resources similar to the leaders. Therefore, we define the third group of forecasters to be located around the 95th percentile of the distribution of the proxy for reputation in the specific market; i.e. on the distribution of the number of analysts employed at the brokerage house for the financial EPS forecasts and on the distribution of the auction houses' turnover for the auctioneers presale estimates. In order to keep the third group comparable to the group of most reputable forecasters, we also select exactly four auction or brokerage houses. These challengers are, from a reputational perspective, close to the reputation leaders and, hence, should have an extra incentive to deviate from the most reputable forecasters' forecasts so as to gain more reputation and to enter the group of reputation leaders in the forecast market.

In the financial forecast market, the brokerage house at the 95th percentile of the *BrokerSize* distribution employs between 31 and 32 analysts.[123] We group the forecast errors of the four brokerage houses in our sample employing 30, 31 and 32 (two houses) analysts and calculate their yearly median forecast errors (*AFEP*).

In the art market, we similarly use the distribution of total sales of the auction houses. The auction houses of the group of reputation leaders are at the top of the distribution with average sales per year (1999-2010) between RMB 798 million and RMB 389 million. The challengers group, defined as the auction houses around the 95th percentile of the distribution, achieves sales per year between RMB 164 million and RMB 104 million. Unfortunately, none of the houses in the third group reports any auctions in the years 2000 and 2001, so we are not able to show a forecast error graph for these years.[124]

Figure 12 and Figure 13 both support Propositions P9 and P10, respectively, and confirm the findings in Figures 10 and 11. In the financial market in Figure 12, the third group of forecasters exhibits a lower yearly median forecast error (*AFEP*) in most of the years than the yearly median forecast error (*AFEP*) of the reputation leaders and of all forecasters combined.

---

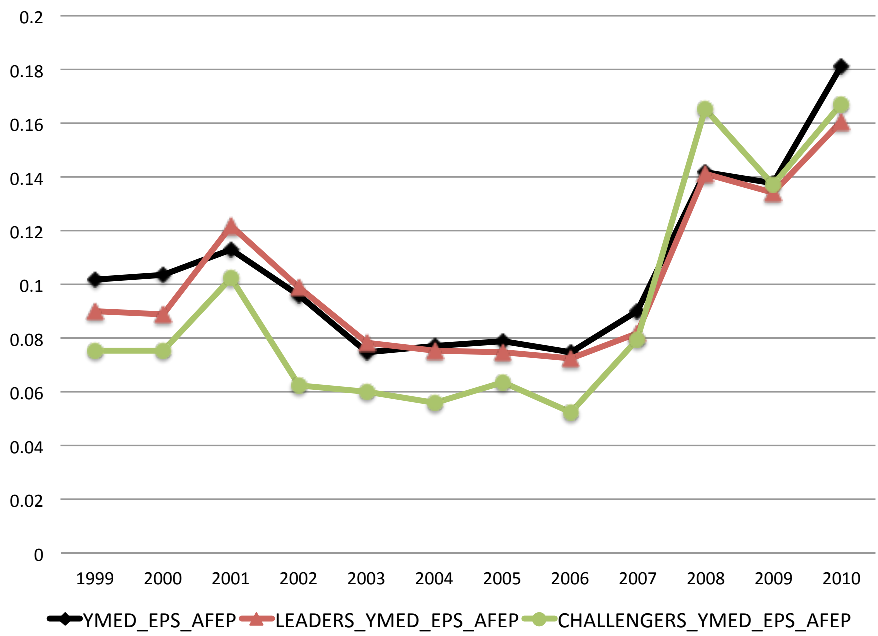
[123] The distribution of brokerage houses on the basis of *BrokerSize* is highly right skewed with a mean of 7.5 analysts and a median of 2 analysts employed. In addition, identifying brokerage houses with 30 analysts is in line with the reputation threshold mentioned in the financial analyst literature lying around 25 to 30 analysts employed (see, e.g., Hong et al. 2000 or Hong and Kubik 2003). Brokerage houses with more than 30 analysts are perceived as belonging to a superior class of brokerage houses.

[124] Also in the art market, the distribution of average sales per year is highly right skewed, with a mean of RMB 34.1 million and a median of RMB 7.5 million.

This supports our argument that in a highly reactive market the reputation leaders induce pressure on the less reputable forecasters to herd with them, as the leaders' forecasts have a high probability of generating a self-fulfilling prophecy. In such a market setting, the third group of forecasters can use this mechanism and issue forecasts with a lower forecast error (AFEP) than the median forecast of the crowd (i.e. of all forecasters).

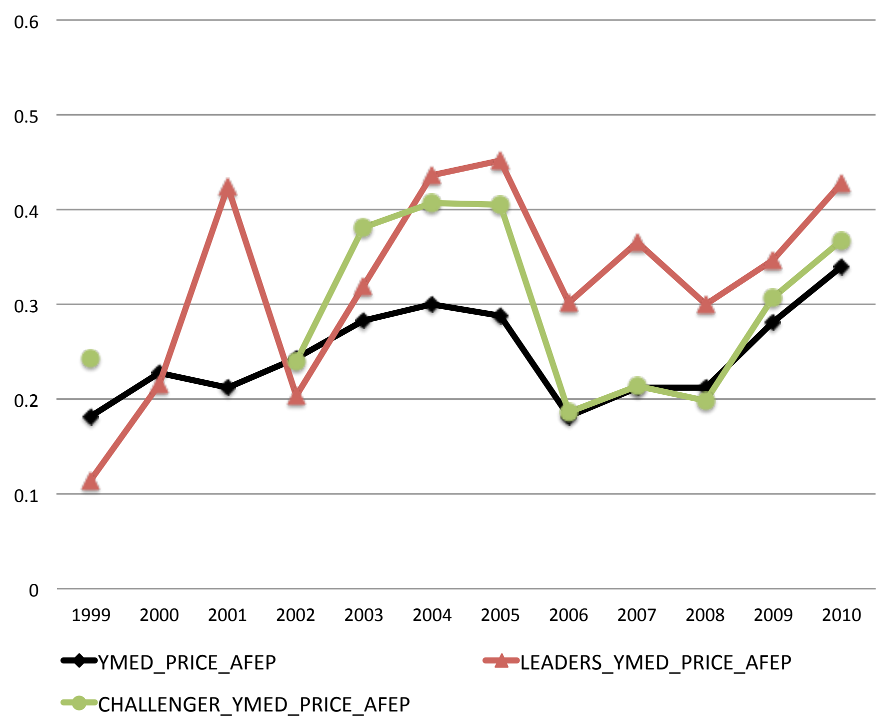
As expected, the art market reveals a different picture than the financial market. Figure 13 presents a yearly median forecast error (*AFEP*) of the challenger group that mostly lies between the median error of the reputation leaders and the median forecast error of all the forecasters combined. This evidence confirms the wisdom-of-crowds argument that the large group of all forecasters in the market will issue more precise forecasts at the median if their forecasts are (sufficiently) independent (i.e. in the absence of herding) due to the low level of reactivity.

Figure 12: Financial Market – Yearly Median Forecast Error (*EPS_AFEP*) of All Forecasters, of the Reputation Leaders and of the Reputation Challengers in the Market for EPS Forecasts



Notes: Yearly median forecast error (*EPS_AFEP*) of all forecasters (*YMED_EPX_AFEP*), of the reputation leaders (*LEADERS_YMED_EPS_AFEP*), and of the reputation challengers (*CHALLENGERS_YMED_EPS_AFEP*) in the market for EPS forecasts; 1999 – 2010. Data source: I/B/E/S

Figure 13: Art Market – Yearly Median Forecast Error (*PRICE_AFEP*) of All Forecasters, of the Reputation Leaders and of the Reputation Challengers in the Market for Art Price Forecasts



Notes: Yearly median forecast error (*PRICE_AFEP*) of all forecasters (*YMED_PRICE_AFEP*), of the reputation leaders (*LEADERS_YMED_PRICE_AFEP*), and of the reputation challengers (*CHALLENGER_ YMED_PRICE_AFEP*) in the market for art price forecasts; 1999 – 2010. Data source: www.artron.net

6.6. Robustness

To provide additional evidence supporting the robustness of our empirical findings for Proposition P1, we employ a time series analysis of the yearly absolute forecast error in percentage of the actual outcome (*AFEP*) for each of the three markets. The idea behind the test is to check if our analysis of the temporary changes in the aggregated forecast error of weather, financial and art forecasts has not fallen prey to a "spurious regression" problem (Granger and Newbold 1974). Such a problem would result in a wrong measurement of the statistical significance in our OLS regressions due to a unit root in the time series. This would mean that the yearly median error of the weather forecasts (*TEMP_AFEP*) would not continuously improve as suggested by the OLS regression. Instead, it would show a spurious correlation between time and the decline of the error.

Therefore we scrutinize our result and use a unit root test (Augmented Dickey-Fuller (ADF) test) for all three time series. Proposition P1 states that the error for weather forecasts

(*TEMP_AFEP*) declines over time, while the errors for art (*PRICE_AFEP*) and financial forecasts (*EPS_AFEP*) do not. Accordingly, we expect the time series of the TEMP_AFEP to be trend stationary, i.e. to have no unit root and to exhibit a clear negative trend, even though there might be periodical oscillations around the trend line. In contrast, the other two markets for art and financial forecasts are expected to show no trend stationarity.

Before we can use the ADF test, we have to define for how many lags we should test the individual time series. Therefore, we employ the improved Akaike Information Criterion (AIC, see Ng and Perron 2001) to identify the number of lags for the test. For all three time series, the AIC suggests the use of 1 lag in the ADF test. Finally, the ADF test supports our first result. The weather forecast error (*TEMP_AFEP*) exhibits a highly significant trend-stationary decline of the error over time (ADF test statistic: -4.67; critical value on the 1% significance level: -4.380). For the time series of the forecast errors in the financial (*EPS_AFEP*) and the art markets (*PRICE_AFEP*), the null hypothesis of unit root cannot be rejected[125]: the *EPS_AFEP* yields a test statistic of 0.063 and a critical value on the 10% significance level of -3.24; the *PRICE_AFEP* yields a test statistic of -2.23 and a critical value at the 10% level of significance of -3.24. We conclude that our finding of a significant decline in weather forecast errors and the findings of an absence of trend in the temporary change of the forecast errors in the art and financial markets are also robust to time-series analysis.

6.7. Concluding Remarks

The hassle of reactivity arises for every forecast that predicts an anthropogenic outcome; the market participants process the new forecast and react according to its information value. By doing so, the forecasts naturally influence the outcome. Hence, the actual outcome is affected by the very forecast that had been issued to predict it. The various problems for forecasting that arise from reactivity are developed in the preceding chapter. In this chapter, we employ a comparative analysis to scrutinize three considerations on forecasts and reactivity from the theoretical part. Firstly, we test the proposition that the forecast error in reactive markets does not improve over time, while it is expected to do so in the non-reactive market for weather forecasts. Secondly, we analyze several propositions regarding the influence of temporal changes in reactivity on the forecast error. Thirdly, we provide evidence for reactivity-

[125] The Augmented Dickey-Fuller test requires that the test statistic be smaller than the critical value at a certain level of significance to reject the null hypothesis of a unit root in a time series (Greene 2002).

induced herding behavior that deteriorates the forecast accuracy of the majority of the forecasters in a highly reactive market.

We employ large datasets in three markets with different levels of reactivity, namely the market for non-reactive weather forecasts, for highly reactive earnings per share forecasts (i.e. financial market), and for weakly reactive presale price forecasts of artworks (i.e. art market).

We find supporting evidence for all propositions. Firstly, a comparison of the time series of the yearly median forecast error in all three markets reveals that the forecast error only diminishes in the non-reactive market for weather forecasts over time; in the other two reactive markets, the forecast error shows no significant trend. Secondly, a comparison between periods of higher and periods of lower reactivity in the same market reveals that reactivity has a substantial influence on the temporal changes in forecast errors. Thirdly, we show that reputation in reactive markets can induce herding behavior through bad learning and that this leads to a distortion of the median forecasts. Such herding behavior results in biased forecast information in highly reactive markets. We are limited in the analysis of herding behavior as we are not able to identify herding directly in the markets due to differences in the forecasting processes and the data available in the different markets. However, we provide evidence that supports the proposition of reactivity-induced herding behavior in a highly reactive market as an obvious explanation of the forecast error pattern of different groups of forecasters that vary in the level of reputation.

Overall, our findings are important in light of the growing yet specious belief that more data and faster data processing will solve all forecasting problems in reactive markets. Additionally, the evidence supports the doubt that the adoption of forecasting techniques that are successful in non-reactive markets (like weather forecasting) will also be successful in reactive markets. We think that this insight is important. It might prevent decision makers from putting too much weight on the capacity of forecasts when using them as sources of policy advice in reactive markets without incorporating the possible reactivity of the market participants to the forecasts issued.

7. Conclusion

This thesis has argued that herding behavior of individuals in organizations and on markets is influenced by the costs that the individuals have to bear if they do not follow the herd. A review of the theoretical and empirical findings in the social-science literature, as well as several anecdotes, reveals that the herding behavior of individuals has detrimental effects on the information aggregation in organizations. If employees face high costs of non-herding, they do not or only sparsely contribute their private information to the pool of knowledge in an organization. This lack of information vitiates the organization's decision-making processes.

In the first part of the thesis (Chapters 2, 3, and 4), the framework of the costs of non-herding is developed and several propositions are deduced and empirically tested. A qualitative study employing in-depth interviews with highly ranked executives, directors and analysts of financial services companies confirms that employees' herding behavior is an important problem for practitioners in general. The study shows that both individual determinants of employees and institutional determinants of their companies influence the employees' costs of non-herding. Furthermore, the responses provide evidence on the asymmetric mechanisms that amplify the detrimental effects of the costs of non-herding: First, they are self-reinforcing. Second and third, they obey an asymmetrical relationship to possible benefits – the costs are both imminent and foreseeable, while the possible benefits both lie in the future and are uncertain. Fourth, an asymmetry exists, as between a private and a common good; the costs have to be borne individually, whereas the benefits are shared with other employees and shareholders of the company.

The subsequent quantitative analysis, employing a large dataset of security analysts' earnings forecasts, reveals an economically and statistically significant influence of one of these institutional determinants, *ownership*, on the analysts' costs of non-herding. Analysts working at large privately held brokerage houses exhibit substantially less herding behavior than similar analysts working at large publicly listed houses. In particular, the study shows that the difference in the costs of non-herding between analysts at privately held and at public brokerage houses is most pronounced for analysts covering financial sector stocks. As security analysts are essential for the provision of information on the future state of financial markets, these findings bear important insights for both investors and regulators.

In the second part of the thesis (Chapters 5 and 6), the reaction of individuals to a certain forecast about a future state of the world is analyzed. This reaction is called *reactivity*. The effects of reactivity on forecasts and opinion formation in several areas of the social and natural sciences are scrutinized. First, reactivity induces the problem that forecasters in reactive systems can never rely on the assessment of their forecast accuracy, as their forecast influences the actual outcome. Second, such first-order problems caused by reactivity produce externalities and biases on all other forecasters at a second-order level. These two orders of problems induced by reactivity are theoretically explained and empirically tested using a comparative analysis with a non-reactive market, namely the market for weather forecasts. It is shown that more data availability and better processing techniques do not solve the problem of reactivity and that the forecast accuracy only improves in non-reactive markets.

Furthermore, the concept of *reactivity-induced herding behavior* is developed: In a reactive market, forecasts induce a self-fulfilling prophecy when enough market participants believe in a certain prediction and start acting according to it. Subsequently, forecasters in the same reactive markets learn about the herding behavior of market participants and understand that such herding behavior leads to a self-fulfilling prophecy for the most reputable forecasters' predictions. Hence, in order to achieve high forecast accuracy, the other forecasters start to follow the forecasters who are perceived to be those with the highest reputation in the market. Finally, such herding behavior among forecasters also exacerbates market participants' herding, because the predictions of the other forecasters are now in line with the most reputable forecaster's predictions. Empirical evidence is provided that supports the proposition of reactivity-induced herding behavior among forecasters. It is shown that, in a highly reactive market, the median forecaster tends to achieve only the same forecast accuracy as the group of most reputable forecasters, suggesting a herding behavior on the part of the median forecaster.[126] In a weakly reactive market, the median forecaster achieves a higher forecast accuracy than the group of most reputable forecasters. This suggests that a weak herding behavior among the bulk of other forecasters has led to a wisdom-of-crowds result (Galton 1907; Surowiecki 2005). The herding mechanism in highly reactive markets is

[126] The median forecaster is calculated as the median of all forecasts other than those of the group of most reputable forecasters.

important, as it distorts the usual reputation process that establishes the quality of goods in a market. In highly reactive markets, reputation loses its disciplining effect on the forecasters, because reputation helps to induce self-fulfilling prophecies which facilitate reactivity-induced herding behavior.

Discussion of results

The measurement of herding behavior is a difficult issue in general. The herding measure employed in Chapter 4 is a simple but dynamic measure which has already been used in several other studies (e.g., Gleason and Lee 2003; Clements and Tse 2005). It accounts for the change of the forecast not only in relation to the herding forecast (i.e. the consensus forecast calculated as the average of all other forecasts) but also in relation to the forecaster's prior forecast. As such, the measure is not static, as earlier measures were (see, e.g., Hong et al. 2000). However, one could think of more complex herding measures that would take into account a larger forecast pattern, i.e. different sequences of new forecast issues, or reactions to other forecasts over a longer time horizon. In addition, one could criticize the fact that the percentage of herding forecasts of about 25% is too low; a "herd" should represent a majority of individuals. Although this is a reasonable objection, it must be said that a group comprising about one quarter of all forecasts can certainly induce herding behavior among market participants. Furthermore, earlier studies that employ a different data sample have about the same proportion of one fourth of herding forecasts (see, e.g., Gleason and Lee 2003; Clements and Tse 2005).

As in most empirical studies, there are causality and self-selection concerns in the empirical analysis of the institutional factor ownership in Chapter 4. The direction of causality, i.e. a privately held brokerage house causes their analysts to exhibit less herding behavior and not vice versa, can be doubted. One can argue that the self-selection of analysts who are less prone to herding behavior into privately held brokerage houses drives this result. There are two arguments against this objection.

Firstly, to identify a correlation between ownership and employees' herding behavior is important above and beyond the question of causality. To know that analysts at larger, privately held brokerage houses engage less in herding behavior is an important message to investors, regulators and other market participants – irrespective of the causal question whether the effect is either due to self-selection or due to causal influence. Even if the analysts self-select into the privately held brokerage houses *only* because they know that they can act more independently in such an environment, the institution of private ownership is still important for information provision in financial markets. Furthermore, the analysis in Chapter 4 provides evidence that supports such a notion. The analysis reveals that the difference in herding behavior between analysts at private and at public brokers is mostly driven by analysts covering financial sector stocks. This is important in two ways: First, analysts that cover financial sector stocks cover their own employer's sector and the stocks of their employer's peer companies. Hence, financial sector analysts working at publicly listed brokerage houses are under strong scrutiny by their supervisors and their management when they release their forecasts; financial sector analysts from private brokers are more independent as their employers are not listed and, thus, their employer's valuation does not depend on their own and other analysts' forecasts. Second, this channel only applies for analysts covering financial sector stocks, because analysts covering stocks from other sectors do not produce such externalities on their own employer's valuation. Hence, we can exclude the possibility that the result is generally driven by self-selection of non-herding analysts into private brokerage houses because of concerns other than the one outlined above. This evidence helps us to exclude self-selection effects other than that analysts covering the financial sector *know* that they can act more independently in privately held brokerage houses and, hence, self-select themselves into them – if they self-select at all. This shows that the institutional factors of brokerage houses and banks are important when it comes to information provision about their peers and the financial market in general and informs investors that they have to select their information sources for this market particularly carefully.

Secondly, one can credibly assume that the analysts' incentive to maximize income is an important factor in choosing their employers. It is a well-known fact that the privately held brokers pay substantially lower wages (see, e.g., Hong and Kubik 2003; Groysberg et al. 2011). The typical career of a successful analyst starts at a smaller, privately held brokerage house of low reputation, then goes on to a more respected, larger private broker, and up to a big and highly prestigious, publicly listed brokerage house (see, e.g., Michaely and Womack 2005; Groysberg et al. 2011). The existence of this career path contradicts the argument that only the self-selection of analysts based on their individual characteristics and preferences – and not the causal effect of the brokers' institution – is responsible for the extent of herding behavior. Although a few analysts might accept a large loss in income because they prefer to work in a privately held brokerage house that allows them to issue more independent forecasts, it seems implausible to conjecture a scenario in which most of the analysts act in such a way. It is more reasonable to assume that on average analysts will work for some time during their careers for larger private brokers before they climb to a high-status position at a big public broker. They use the time and the environment at larger private brokers to issue independent, non-herding forecasts in order to present themselves and their ability to the market.

When considering the second part of the thesis, the evidence provided on the interference between reactive market outcomes and forecasts is by necessity limited. Reactivity is, by definition, difficult to tackle with empirical methods. Ideally, one would prefer an identical situation to occur twice, once with and once without a forecast being issued. Due to the lack of such perfect settings and the endogeneity problem between forecasts and reactive outcomes, this thesis employs a simple empirical approach.[127] It compares the development of forecast accuracy over time between three different markets and the underlying changes in market fundamentals and sentiments, which influence the level of reactivity in a market. Additionally, it scrutinizes the difference between various groups of forecast providers in the financial and art markets. However, such an analysis could be extended by developing techniques for reactive markets which would be able to reconstruct a situation with forecasts as if they had not been issued. Although such attempts have been made in the financial (e.g., Simpson, forthcoming) and art markets (Mei and Moses 2005), an approach that delivers comparable results in both markets by employing the same procedure still has to be undertaken.

Implications for economics

This thesis leads to several implications for economics. First, it contributes to the economics literature by providing robust evidence that institutional factors of organizations influence employees' herding behavior. The framework of the costs of non-herding offers a theoretical basis on which to analyze determinants of employees' herding behavior. The interview study reveals that the framework is useful for examining various factors which explain why employees do not differ from the prevailing opinion and how asymmetric incentives can

[127] Another way would be to conduct a laboratory experiment, in which the various situations are reproducible, but this approach is confronted with a major problem of external validity. The study aims at eliciting the existence of reactivity in a real-life setting, employing recorded data from typical market situations.

explain a self-reinforcing dynamic of employees' herding behavior in organizations. Additionally, the study identifies several relevant institutional factors that shape the employees' costs of non-herding: the hierarchical structure, the degree of decentralization in an organization, the ownership structure, the degree of performance pay, and the independence of the internal risk management.

Second, based on the framework and supporting the qualitative evidence, the quantitative study shows that the ownership structure of brokerage houses and banks has an economically and statistically significant effect on the herding behavior of their employees. The evidence provided expands the literature on security analysts in particular, and the empirical literature on employees' herding behavior in general.

Third, reactivity and its effects on the ex-post assessment of forecasts influence research in areas of the social sciences that focus on forecasters' and other information providers' herding behavior. As market participants react to forecasts and this reaction changes the actual outcome, the impossibility of an unbiased ex-post analysis of the forecasts' accuracy hinders the sound development of forecasting theory and its empirical testing. This thesis is not the first treatise on the problems of reactivity and the possibility of self-fulfilling or self-defeating prophecies (see, e.g., Morgenstern 1928; Merton 1936). Yet, it reminds us of the difficulties of economic research induced by the reactive behavior of market participants. It also links the various areas of the social sciences where reactivity exists. Economic research might benefit from the literature review, which draws extensively on works from sociology and political science in addition to the economic literature. Several fields of application are discussed, reaching from university rankings to Internet-based forecasting with Google trends. As discussed in Chapters 5 and 6, the problems for forecasting induced by reactivity are not or only partially taken into account in economics. In particular, empirical research analyzing the performance and accuracy of forecasting in economics and in financial markets largely neglects the distorting effects of reactivity. For example, in the literature on the forecast accuracy of security analysts, Beyer (2008) was the first to theoretically describe the problems for analysts that are induced by the reactive behavior of managers, who try to meet a certain earnings forecast. Yet even today most empirical papers (see, e.g., Cohen et al. 2010; Clement et al. 2011; Bonini and Kerl 2012) do not take this problem into account (Liu and Natarajan 2012 or Bissessur and Veenman 2013, being notable exceptions).

Fourth, this thesis proposes a mechanism of *reactivity-induced* herding behavior in Chapter 5, adding a relevant extension to the literature on herding behavior in economics and the social sciences in general. The empirical study in Chapter 6 provides evidence that a high level of reactivity in a market can induce herding behavior among forecasters. Herding is measured as an aggregated, congruent development of forecast accuracy between the different groups of forecasters. The mechanism is distinctively different from the usual herding theories proposed in the literature, such as reputational or informational herding behavior (see, e.g., Bhikhchandani and Sharma 2001; Hirshleifer and Teoh 2003; 2009). The argument goes above and beyond the problem that less able forecasters could mimic more able forecasters, as in the case of reputation-based herding behavior (Scharfstein and Stein 1990; Zwiebel 1995). Rather, it is stated that reactivity changes the outcome such that a wrong prediction by a less able but ex-ante reputable forecaster may become reality and, thus, ex-post be perceived as a correct prediction. Likewise, the argument here is not based on an informational cascade (Banerjee 1992; Bikhchandani et al. 1992). The herding forecasters might even be aware that they issue mistaken forecasts, but because the actual outcome is strongly influenced by reactivity (leading to a self-fulfilling prophecy), they still engage in herding behavior. All the subjects act in a fully rational manner.

Implications for practitioners

For practitioners, the implications of this thesis are twofold. Firstly, the interview study in Chapter 3 reveals that managers of banks and financial services providers are aware of the problems posed by herding behavior. However, they seem not to actively engage in reducing the costs of non-herding. In addition, managers of larger companies are aware of suggested internal rules of conduct in meetings, but state in the interviews that these rules are not used in daily life. Based on the managers' responses, this thesis reminds us that such procedures can help to lower the employees' costs of non-herding and to improve the information pool available for the organization's decision-making process.

Secondly, the thesis primarily focuses on institutional factors at the organizational macro level. The interview study identifies four such factors that help to lower the employees' costs of non-herding: a flatter hierarchy with a more decentralized organization; privately held companies; a lower portion of performance pay; and a more independent risk management. Managers and entrepreneurs should consider these factors when organizing or establishing a firm. The institutional factors at the macro level go above and beyond the usual aspects mentioned in the literature on herding behavior (e.g., Morrison and Milliken 2000; 2003; Hirshleifer and Teoh 2003; 2009) and in the practitioners' literature that mostly focuses on institutional factors at the micro level, like, e.g., *Robert's Rules of Order* (Robert 1876 [2011]; Shiller 1995, p. 183) or the guides based on Janis's (1972) ideas to restrain groupthink (e.g., Topchick 2007; Wilcox 2010).

Two interesting institutional factors at the micro level are identified which have not been found in the practitioners' literature. Managers propose more informal communication before and between meetings, and want employees to get to know each other across different units and ranks of an organization to help lower the employees' costs of non-herding. They explain that these two factors could help share ideas that challenge the prevailing opinion in a firm at lower costs of non-herding.

Implications for public policy

The empirical study in Chapter 4 aims at informing investors and regulators about the importance of institutional factors of information providers in financial markets. Analysts covering financial sector stocks and working with larger privately held brokerage houses exhibit a substantially lower propensity to engage in forecast herding than identical analysts working with publicly listed brokers and banks. The last financial crisis provides many examples of analysts who were working with large and prestigious, publicly listed brokerage houses and did not release information about the imminence of the crisis.

The study argues that financial sector analysts at public brokers cannot independently reveal their private information as their forecasts influence the valuation of their own employer, whose stocks are listed in the same sector. Due to the externality of their forecasts on their own employer, financial sector analysts at public brokers cannot reveal their private information to the market. Instead, they are scrutinized and restrained by their own management when trying to issue critical information about their own sector. Such an internal forecast policy creates incentives for financial sector analysts at major listed brokers and banks to engage in herding behavior to avoid facing the costs of non-herding when going against the opinion of their own management. Such distorted incentives create a lack of independent information provision in financial markets.

Since the second most important source of information for financial markets, the credit rating agencies, depend on business with such big banks and brokers, these two important information providers form a kind of unholy alliance (see also Mathis et al. 2009; Fong et al. 2012). Hence, regulators should be interested in removing the interdependence of these two groups and in fostering a more independent information provision in financial markets. This might be achieved by supporting larger privately held brokerage houses or by barring the provision of financial sector forecasts from analysts working at publicly listed banks and brokerage houses. Investors and customers can also reduce the problem of herding behavior in financial markets by carefully selecting the source of financial sector forecasts and by refraining from reacting too credulously to forecasts from analysts of publicly listed banks or brokers.

It has often been argued that the reputational concerns of forecast providers help to restore the incentives to deliver accurate forecasts, as consumers choose the forecaster whose prediction has been most accurate in the preceding period.[128] It is conjectured in Chapter 5 and supported partially by evidence in Chapter 6 that such a mechanism does not solve the incentive problem for forecasters in reactive markets – in contrast, it may even have a detrimental effect on forecasters' incentives. It is argued that the reactivity of the market participants, who use the forecasts and produce the market outcome (i.e. the price of a stock or another security) by reacting to the forecast, leads to distorted forecasters' incentives. Once a forecast provider makes a highly accurate forecast in a certain period, the market participants might attribute a high reputation to this forecaster and react to his or her forecasts such that a self-fulfilling prophecy occurs. Such reactivity potentially leads to a high accuracy of the forecast issued by the same forecaster, after he or she has become highly reputed. Such a result not only distorts the incentive effect of reputation but induces a detrimental herding behavior among other forecasters. The other forecasters then engage in reactivity-induced herding behavior, because they anticipate that the market participants' reactions will probably again lead to a self-fulfilling prophecy and create a highly accurate forecast. In addition, the increasingly concentrated forecasts in the market as the forecasters herd provide a distorted signal of general agreement on the future to the market, making the market participants even more likely to follow the herd of forecasters.

In systems where forecasts do not influence the actual outcome, reactivity cannot occur. Hence in Chapter 6, a comparative institutional analysis between the non-reactive market for weather forecasts and the reactive art and financial markets revealed that the forecast accuracy in reactive markets has not improved over the last decade. Only the forecast accuracy regarding the non-reactive weather system has made progress. This finding is even more

[128] See Mathis et al. 2009 for a critical analysis on the effect of reputation in the market for credit rating agencies.

interesting when it is recalled that the dramatic increase in the amount of data and the data processing took place in all three markets. Hence, it suggests that the transfer of forecasting techniques from non-reactive to reactive markets will not provide the expected improvement in forecast accuracy. This stands in contrast, for example, to a recent paper by the Bank of International Settlement (Vahey and Wakerly 2013) and in other literature on economic forecasting (e.g., Garratt et al. 2011).

Reactivity develops in every anthropogenic market, where people react to forecasts and create the final outcome of the market. Chapter 5 proposes the institution of the devil's advocate, which would have to provide counterarguments and forecasts against the most reputable forecast provider(s). The proposition developed in this thesis is radical, as it clearly assigns the function of arguing against any most reputable forecaster or prevailing opinion in the market to the institution of the devil's advocate. To fulfill such a role, it is important that these boards are independent of any private or public bodies. Hence, examples of publicly funded institutions which issue economic forecasts, such as the Council for Economic Advisers in the U.S. or the German Council of Economic Experts, do not or only partially satisfy these requirements. Chapter 5 also explains why many international organizations do not fulfill the role of a devil's advocate. They usually incorporate not only units that generate forecasts but also other units that have to react officially to the forecasts when they plan and execute their duties. For example, the International Monetary Fund generates economic forecasts for almost every country in the world but also plans and executes its own financial support in accordance with the forecasts it releases. In relation to the organization's market power and the size of the countries' economies, these forecasts potentially trigger a reactivity-induced herding behavior among other forecasters and market participants in anticipating the actions of the organization according to its forecasts. Thus, it is important that the institution of the devil's advocate *only* produces forecasts and that it is independent of any executive body or organization that has to use and act upon forecasts. The German Council of Economic Experts, for example, fulfills this important aspect. Yet, its members are elected by the Federal President and not independent of the government.

Prospects for future research

The analysis conducted in this thesis provides insights into the phenomenon of herding behavior and reactivity in organizations and markets but inevitably leaves many questions unaddressed. Although the interview study touches on many different determinants of the costs of non-herding in organizations, the quantitative study focuses mainly on the institutional factor of ownership while controlling for other organizational and individual parameters. Many other institutional factors revealed by the interview study in Chapter 3 should be further quantitatively investigated and their interactions should be studied. For example, the influence of the role of risk management or the degree of performance pay should be scrutinized both individually and in combination with other institutional factors.

In particular, it would be interesting to ascertain how the individually faced costs of non-herding aggregate to a herd behavior of firms and other organizations in specific markets. High individual costs of non-herding engender a strong and cohesive prevailing opinion in an organization. This prevailing opinion inside a firm could induce herding behavior among firms in a specific market, depending on the firm's market share. Finally, this could lead to a strongly biased perception of opinions and predictions in markets. Firms might issue similar forecasts about a future state of the world although in each firm only a small majority of employees (or even a minority) share this forecast and an almost equally large portion of employees contradict it. So it seems important to further analyze the aggregation of opinions in markets and how a higher concentration of market shares leads to a stronger bias of the available information in a certain market (see Hong and Kacperczyk 2010). Future research should examine such interdependencies between individual and aggregated herd behavior.

In addition, future research on social learning and herding behavior should consider reactivity as an important part of a more extensive explanation of herding behavior in organizations and markets. Furthermore, the institution of the devil's advocate proposed to reduce the reactive behavior of market participants has to be scrutinized empirically and developed further.

The evidence that forecast accuracy in reactive markets has not improved over the last decade despite the technological progress in generating and processing big data leads to the conclusion that the problem of reactivity cannot be tackled by improving forecasting techniques *per se*. Sooner or later, the market participants' reactions will erase the improvement. In contrast, a better balancing of the various incentives of forecasters might help to improve information provision in reactive markets. This thesis proposes institutional improvements to incorporate more balanced incentives in organizations and markets to provide a richer informational environment for all market participants.

References

Abarbanell, J., Lehavy, R., 2003. Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts. *Journal of Accounting and Economics* 36: 105-146.

Acemoglu, D., Johnson, S., Robinson, J., 2005. The rise of Europe: Atlantic trade, institutional change, and economic growth. *The American Economic Review* 95: 546-579.

Agrawal, A., Chen, M.A., 2008. Do Analyst Conflicts Matter? Evidence from Stock Recommendations. *Journal of Law and Economics* 51: 503-537.

Akerlof, G.A., Shiller, R.J., 2009. *Animal Spirits*. Princeton: Princeton University Press.

Alevy, J.E., Haigh, M.S., List, J.A., 2007. Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance* 62: 151-180.

American, Meteorological, Society, 2008. Enhancing Weather Information with Probability Forecasts. *Bulletin of the American Meteorological Society* 89: 1-12.

Argyris, C., Schön, D., 1978. *Organizational Learning*. Reading: Addison-Wesley.

Arthur, W.B., 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The Economic Journal* 99: 116-131.

ArtPrice, 2011. *Art Market Trends 2011*. Available online at http://web.artprice.com/artmarketinsight (accessed on April 21, 2013).

Asch, S.E., 1951. *Effects of Group Pressure upon the Modification and Distortion of Judgements*. In: Guetzkow M. (ed.). Groups, Leadership, and Men. Pittsburgh: Carnegie Press.

Asch, S.E., 1956. Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs: General and Applied* 70: 1-70.

Ashenfelter, O., 1989. How auctions work for wine and art. *The Journal of Economic Perspectives* 3: 23-36.

Ashenfelter, O., Graddy, K., 2003. Auctions and the Price of Art. *Journal of Economic Literature* 41: 763-787.

Ashenfelter, O., Graddy, K., 2006. *Art auctions*. London: Elsevier.

Ashenfelter, O., Graddy, K., 2011. Sale Rates and Price Movements in Art Auctions. *The American Economic Review* 101: 212-216.

Ashiya, M., Doi, T., 2001. Herd behavior of Japanese economists. *Journal of Economic Behavior & Organization* 46: 343-346.

Aspara, J., Pajunen, K., Tikkanen, H., Tainio, R., 2008. Explaining Sacrifice of Long-Term Investments for Short-Term Earnings: Spiral of Silence in Financial Markets. *Proceedings of the Annual Meeting of the Academy of Management*

Avery, C., Zemsky, P., 1998. Multidimensional Uncertainty and Herd Behavior in Financial Markets. *The American Economic Review* 88: 724-748.

Bai, J., Guo, J., Mandel, B., 2013. Going global: markups and product quality in the Chinese art market. *FRB of New York Staff Report.*

Bai, J., Philippon, T., Savov, A., 2012. Have financial markets become more informative? *FRB of New York Staff Report.*

Baker, M., Wurgler, J., 2006. Investor Sentiment and the Cross Section of Stock Returns. *The Journal of Finance* 61: 1645-1680.

Baker, M., Wurgler, J., 2007. Investor Sentiment in the Stock Market. *Journal of Economic Perspectives* 21: 129-152.

Bandura, A., 1965. Influence of Model's Reinforcement Contingencies on the Aquisition of Imitative Responses. *Journal of Personality and Social Psychology*: 589-595.

Bandura, A., 1976. *Social Learning Theory*. In: Spence J.T., Carson R.C. & Thiebaut J.W. (eds.). Behavioral Approaches to Therapy. Morristown NJ: General Learning Press.

Banerjee, A., Somanathan, R., 2001. A Simple Model of Voice. *Quarterly Journal of Economics* 116: 189-227.

Banerjee, A.V., 1992. A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107: 797-817.

Bantel, K.A., Jackson, S.E., 1989. Top Management and Innovations in Banking: Does the Composition of the Top Team Make a Difference? *Strategic Management Journal* 10: 107-124.

Bao, J., Edmans, A., 2011. Do Investment Banks Matter for M&A Returns? *Review of Financial Studies* 24: 2286-2315.

Barber, B.M., Lehavy, R., Trueman, B., 2007. Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics* 85: 490-517.

Baron, R.S., 2005. *So Right It's Wrong: Groupthink and the Ubiquitous Nature of Polarized Group Decision Making*. In: Zanna M.P. (ed.). Advances in experimental social psychology. San Diego: Elsevier Academic Press. 219-253.

Baumol, W.J., 1957. Interactions between successive polling results and voting intentions. *Public Opinion Quarterly* 21: 318-323.

Beatty, A., Harris, D.G., 1999. The effects of taxes, agency costs and information asymmetry on earnings management: A comparison of public and private firms. *Review of Accounting Studies* 4: 299-326.

Beatty, A.L., Ke, B., Petroni, K.R., 2002. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* 77: 547-570.

Beggs, A., Graddy, K., 1997. Declining values and the afternoon effect: Evidence from art auctions. *The Rand Journal of Economics*: 544-565.

Beggs, A., Graddy, K., 2009. Anchoring effects: Evidence from art auctions. *The American Economic Review* 99: 1027-1039.

Bénabou, R., 2013. Groupthink: Collective delusions in organizations and markets. *The Review of Economic Studies* 80: 429-462.

Benz, M., Frey, B.S., 2008. The value of doing what you like: Evidence from the self-employed in 23 countries. *Journal of Economic Behavior & Organization* 68: 445-455.

Berger, A.N., Miller, N.H., Petersen, M.A., Rajan, R.G., Stein, J.C., 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics* 76: 237-269.

Bernhardt, D., Campello, M., Kutsoati, E., 2006. Who herds? *Journal of Financial Economics* 80: 657-675.

Besley, T., Case, A., 2003. Political institutions and policy choices: evidence from the United States. *Journal of Economic Literature* 41: 7-73.

Beunza, D., Stark, D., 2012. From dissonance to resonance: cognitive interdependence in quantitative finance. *Economy and Society* 41: 383-417.

Bewley, R., Fiebig, D.G., 2002. On the herding instinct of interest rate forecasters. *Empirical Economics* 27: 403-425.

Bewley, T.F., 1998. Why not cut pay? *European Economic Review* 42: 459-490.

Bewley, T.F., 1999. *Why Wages Don't Fall During A Recession*. Cambridge, MA: Harvard University Press.

Beyer, A., 2008. Financial analysts, forecast revisions and managers reporting behavior. *Journal of Accounting and Economics* 46: 334-348.

Beyer, A., Cohen, D.A., Lys, T.Z., Walther, B.R., 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50: 296-343.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *The Journal of Political Economy* 100: 992-1026.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1998. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *The Journal of Economic Perspectives* 12: 151-170.

Bikhchandani, S., Sharma, S., 2001. Herd Behavior in Financial Markets. *IMF Staff Papers* 47: 279-310.

Binswanger, D., Gimes, M., 2011. Wir Brauchen schaerfere Regeln. *Das Magazin*, May 13, 2011: 16-21.

Bissessur, S., Veenman, D., 2013. The Role of Ex-Ante Uncertainty in Explaining Why Firms Meet or Just Beat Analysts' Earnings Forecasts. *Working Paper - INSEAD.*

Blinder, A.S., Choi, D.H., 1990. A shred of evidence on theories of wage stickiness. *The Quarterly Journal of Economics* 105: 1003-1015.

Bloom, N., Van Reenen, J., 2007. Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics* 122: 1351-1408.

Bloom, N., Van Reenen, J., 2010. Why do management practices differ across firms and countries? *Journal of Economic Perspectives* 24: 203-224.

Boll, D., 2011. Boom im Reich der Mitte. *NZZ am Sonntag*, April 3, 2011: 51.

Bolton, P., Dewatripont, M., 2005. *Contract theory*. Cambridge: MIT Press.

Bombach, G., 1962. Ueber die Moeglichkeit wirtschaftlicher Voraussagen. *Kyklos* 15: 29-67.

Bonini, S., Kerl, A., 2012. Private information, subjective valuation and target price accuracy. *Subjective Valuation and Target Price Accuracy (October 30, 2012)*

Bowen, F., Blackmon, K., 2003. Spirals of Silence: The Dynamic Effects of Diversity on Organizational Voice. *Journal of Management Studies* 40: 1393-1417.

Bowles, S., 1998. Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature* 36: 75-111.

Buizza, R., Hagedorn, R., Isaksen, L., 2010. *Recent changes of the ECMWF Ensemble Prediction System*. In: European Center for Medium-Range Weather Forecasts Conference.Vitart F. (Ed.). Zurich.

Burgstahler, D., Eames, M., 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting* 33: 633-652.

Burt, R.S., 1982. *Toward a Structural Theory of Action: Network Models of Social Structure, Perception and Action*. London, NJ: Academic Press.

Butler, D., 2008. Web data predict flu. *Nature* 456: 287.

Butler, D., 2013. When Google got flu wrong. *Nature* 494: 155.

Callon, M., 1998. *The laws of the markets*. Oxford: Blackwell.

Campbell, D.T., 1957. Factors relevant to the validity of experiments in social settings. *Psychological Bulletin* 54: 297-319.

Campbell, D.T., Stanley, J.C., 1963. *Experimental and quasi-experimental designs for research on teaching*. New Haven: Ceneage Learning.

Chamley, C., 2004. *Rational herds: Economic models of social learning*. Cambridge: Cambridge University Press.

Chang, X., Dasgupta, S., Hilary, G., 2006. Analyst coverage and financing decisions. *The Journal of Finance* 61: 3009-3048.

Chen, Q., Jiang, W., 2006. Analysts weighting of private and public information. *Review of Financial Studies* 19: 319-355.

Cheong, F.S., Thomas, J., 2011. Why do EPS forecast error and dispersion not vary with scale? Implications for analyst and managerial behavior. *Journal of Accounting Research* 49: 359-401.

Chevalier, J., Ellison, G., 1999. Career Concerns of Mutual Fund Managers. *Quarterly Journal of Economics* 114: 389-432.

Choi, H., Varian, H., 2012. Predicting the present with google trends. *Economic Record* 88: 2-9.

Choi, N., Sias, R.W., 2009. Institutional industry herding. *Journal of Financial Economics* 94: 469-491.

Christie, W.G., Huang, R.D., 1995. Following the Pied Piper: Do Individual Returns Herd around the Market? *Financial Analysts Journal* 51: 31-37.

Chung, R., 2012. *The Freedom to Publish Opinion Poll Report*. World Association of Public Opinion Research Research (Ed.). University of Hong Kong, Gallup, WAPOR.

Cipriani, M., Guarino, A., 2005. Herd Behavior in a Laboratory Financial Market. *The American Economic Review* 95: 1427-1443.

Cipriani, M., Guarino, A., 2009. Herd behavior in financial markets: an experiment with financial market professionals. *Journal of the European Economic Association* 7: 206-233.

Clement, M.B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285-303.

Clement, M.B., Hales, J., Xue, Y., 2011. Understanding analysts' use of stock returns and other analysts' revisions when forecasting earnings. *Journal of Accounting and Economics* 51: 279-299.

Clement, M.B., Tse, S.Y., 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78: 227-249.

Clement, M.B., Tse, S.Y., 2005. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance* 60: 307-341.

Clements, M.P., Hendry, D.F., 2011. *The Oxford handbook of economic forecasting*: Oxford University Press.

Cohen, D.A., Dey, A., Lys, T.Z., 2008. Real and accrual-based earnings management in the pre-and post-Sarbanes-Oxley periods. *The Accounting Review* 83: 757-787.

Cohen, L., Frazzini, A., Malloy, C., 2010. Sell-Side School Ties. *The Journal of Finance* 65: 1409-1437.

Corwin, S.A., Schultz, P., 2005. The role of IPO underwriting syndicates: Pricing, information production, and underwriter competition. *The Journal of Finance* 60: 443-486.

Cowen, A., Groysberg, B., Healy, P., 2006. Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41: 119-146.

Credit Suisse, 2007. Annual Report.

Cueni, R., Fiechter, P., 2013. Bold Analysts or Bold Banks: Do Institutional Factors Drive Herding Behavior? *Working Paper (SSRN).*

Cueni, R., Frey, B.S., 2012. *How to Overcome Herding Behavior in Firms*. In: Towards a New Theory of the Firm. Rosannas J.M. & Ricart J.E. (Eds.). Bilbao: Fundacion BBVA.

Cueni, R., Frey, B.S., 2013. Forecasts and Reactivity. *mimeo, University of Zurich.*

Cueni, R., Frey, B.S., forthcoming. *Repressed Voice and Costs of Non-Herding*. In: Contract Governance - Dimensions in Law and Interdisciplinary Research. Grundmann S., Moeslein F. & Riesenhuber K. (Eds.). Oxford: Oxford University Press.

Daniel, K., Hirshleifer, D., Teoh, S.H., 2002. Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics* 49: 139-209.

David, P.A., 1985. Clio and the Economics of QWERTY. *The American Economic Review* 75: 332-337.

Degeorge, F., Patel, J., Zeckhauser, R., 1999. Earnings Management to Exceed Thresholds. *The Journal of Business* 72: 1-33.

Deutsch, M., Gerard, H.B., 1955. A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology* 51: 629.

Dewatripont, M., Tirole, J., 1999. Advocates. *Journal of Political Economy* 107: 1-39.

Diekmann, A., 2001. *Empirische Sozialforschung*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.

Dimson, E., Spaenjers, C., 2011. Ex post: The investment performance of collectible stamps. *Journal of Financial Economics* 100: 443-458.

Drehmann, M., Oechssler, J., Roider, A., 2005. Herding and Contrarian Behavior in Financial Markets: An Internet Experiment. *The American Economic Review* 95: 1403-1426.

Dyck, A., Morse, A., Zingales, L., 2010. Who Blows the Whistle on Corporate Fraud. *The Journal of Finance* 65: 2213-2253.

Ekelund Jr, R.B., Ressler, R.W., Watson, J.K., 1998. Estimates, bias and sales in Latin-American art auctions, 1977-1996. *Journal of Cultural Economics* 22: 33-42.

Elliott, G., Timmermann, A., 2008. Economic forecasting. *Journal of Economic Literature*: 3-56.

Ely, R.J., 2004. A Field Study of Group Diversity, Participation in Diversity Education Programs, and Performance. *Journal of Organizational Behavior* 25: 755-780.

Esman, A., 2012. China's USD 13 Billion Art Fraud - and What it Means For You. *Forbes*, August 13, 2012.

Espeland, W.N., Sauder, M., 2007. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113: 1-40.

Esposito, E., 2013. The structures of uncertainty: performativity and unpredictability in economic operations. *Economy and Society* 42: 102-129.

Ettredge, M., Gerdes, J.J., 2005. Timeliness of investor relations data at corporate web sites. *Communications of the ACM* 48: 95-100.

Evans, G.W., Honkapohja, S., 2001. *Learning and expectations in macroeconomics*. Princeton: Princeton University Press.

Eysenbach, G., 2006. *Infodemiology: tracking flu-related searches on the web for syndromic surveillance*. In: AMIA Annual Symposium Proceedings: 244, American Medical Informatics Association.

Feasby, C.C., 1997. Public Opinion Poll Restrictions, Elections, and the Charter. *University of Toronto – Faculty of Law Review* 55: 241-268.

Fernando, C.S., May, A.D., Megginson, W.L., 2012. The value of investment banking relationships: evidence from the collapse of Lehman Brothers. *The Journal of Finance* 67: 235-270.

Festinger, L., 1954. A Theory of Social Comparison Processes. *Human Relations* 7: 114-140.

Festinger, L., 1957. *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Fong, K., Hong, H., Kacperczyk, M., Kubik, J., 2012. Do Security Analysts Discipline Credit Rating Agencies? *AFA 2013 San Diego Meetings Paper*.

Francis, J., Philbrick, D., 1993. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* 31: 216-230.

Frank, R.H., 1999. Higher education: The ultimate winner-take-all market? *Cornell Working Paper Series (CHERI).*

Frank, R.H., Cook, P.J., 1988. *The Winner-Take-All Society: How More and More Americans Compete for Ever Fewer and Bigger Prizes, Encouraging Economic Waste, Income Inequality, and an Impoverished Cultural Life*. New York: The Free Press.

Frey, B.S., 2008. *Happiness: A Revolution in Economics*. Cambridge: MIT Press.

Frey, B.S., Gallus, J., 2012. Happiness policy and economic development. *International Journal of Happiness and Development* 1: 102-111.

Frey, B.S., Stutzer, A., 2002a. What can economists learn from happiness research? *Journal of Economic Literature* 40: 402-435.

Frey, B.S., Stutzer, A., 2002b. *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton: Princeton University Press.

Frey, B.S., Stutzer, A., 2009. *Should happiness be maximized?* In: Dutt A.K. & Radcliff B. (eds.). Happiness, Economics and Politics. Cheltenham: Edward Elgar. 97-126.

Galton, F., 1907. Vox populi. *Nature* 75: 450-451.

Garber, P.M., 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. Cambridge: MIT Press.

Garcia-Ferrer, A., 2012. On Granger's predictability of financial markets in theory and practice. *International Journal of Forecasting* 28: 121-127.

Garicano, L., Posner, R.A., 2005. Intelligence Failures: An Organizational Economics Perspective. *The Journal of Economic Perspectives* 19: 151-170.

Garratt, A., Mitchell, J., Vahey, S.P., Wakerly, E.C., 2011. Real-time inflation forecast densities from ensemble Phillips curves. *The North American Journal of Economics and Finance* 22: 77-87.

Gentzkow, M., Shapiro, J.M., 2006. Media Bias and Reputation. *Journal of Political Economy* 114: 280-316.

Giles, J., 2011. Social science lines up its biggest challenges. *Nature* 470: 18-19.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L., 2009. Detecting influenza epidemics using search engine query data. *Nature* 457: 1012-1014.

Glaeser, J., Laudel, G., 2006. *Experteninterviews und qualitative Inhaltsanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaft.

Gleason, C.A., Lee, C.M.C., 2003. Analyst forecast revisions and market price discovery. *Accounting Review*: 193-225.

Goodhart, C., 1975. Problems of monetary management: the UK experience. *Papers in Monetary Economics (Reserve Bank of Australia)* 26: 219-245.

Graham, J.R., 1999. Herding among Investment Newsletters: Theory and Evidence. *The Journal of Finance* 54: 237-268.

Graham, J.R., Harvey, C.R., Rajgopal, S., 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3-73.

Granger, C.W., Newbold, P., 1974. Spurious regressions in econometrics. *Journal of Econometrics* 2: 111-120.

Granovetter, M.S., 1973. The Strength of Weak Ties. *The American Journal of Sociology* 78: 1360-1380.

Greene, W.H., 2003. *Econometric Analysis*. Massachusetts: Pearson Education

Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *The American Economic Review* 85: 1088-1105.

Grossman, S.J., 1977. The existence of futures markets, noisy rational expectations and informational externalities. *The Review of Economic Studies* 44: 431-449.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. *The American Economic Review* 70: 393-408.

Groysberg, B., Healy, P.M., Maber, D.A., 2011. What Drives Sell-Side Analyst Compensation at High-Status Investment Banks? *Journal of Accounting Research* 49: 969-1000.

Grunberg, E., Modigliani, F., 1954. The predictability of social events. *The Journal of Political Economy* 62: 465-478.

Grunberg, E., Modigliani, F., 1963. Economic forecasting when the subject of the forecast is influenced by the forecast: comment. *The American Economic Review* 53: 734-737.

Guadalupe, M., Li, H., Wulf, J., 2012. Who lives in the c-suite? Organizational structure and the division of labor in top management. *NBER Working paper.*

Hamilton, J.D., 2009. Causes and Consequences of the Oil Shock of 2007-08. *NBER Working paper.*

Hamilton, W.D., 1971. Geometry for the selfish herd. *Journal of Theoretical Biology* 31: 295-311.

Hands, D.W., 1990. Grunberg and Modigliani, Public Predictions, and the New Classical Macroeconomics. *Research in the History of Economic Thought and Methodology* 7: 207-223.

Hayward, M.L., Boeker, W., 1998. Power and conflicts of interest in professional firms: Evidence from investment banking. *Administrative Science Quarterly*: 1-22.

Healy, P.M., Palepu, K.G., 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31: 405-440.

Healy, P.M., Wahlen, J.M., 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13: 365-383.

Heilemann, U., Stekler, H.O., 2012. Has The Accuracy of Macroeconomic Forecasts for Germany Improved? *German Economic Review*. 14: 235-253.

Heisenberg, W., 1927. Ueber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift fuer Physik* 43: 172-198.

Helper, S., 2000. Economists and Field Research:" You Can Observe a Lot Just by Watching". *The American Economic Review* 90: 228-232.

Henshel, R.L., 1993. Do self-fulfilling prophecies improve or degrade predictive accuracy? How sociology and economics can disagree and both be right. *The Journal of Socio-Economics* 22: 85-104.

Henshel, R.L., 1995. The Gruenberg/Modigliani and Simon possibility theorem: A social psychological critique. *Journal of Socio-Economics* 24: 501-520.

Hertzberg, A., Liberti, J.M., Paravasini, D., 2010. Information and Incentives Inside the Firm: Evidence from Loan Office Rotation. *Journal of Finance* 65: 795-828.

Hey, J.D., Morone, A., 2004. Do Markets Drive Out Lemmings,or Vice Versa? *Economica* 71: 637-659.

Hilary, G., Hsu, C., 2013. Analyst Forecast Consistency. *Journal of Finance* 1: 271–297.

Hirschman, A.O., 1970. *Exit, Voice and Loyalty - Responses to Decline in Firms, Organizations, and States*. Cambridge: Harvard University Press.

Hirshleifer, D., Hong Teoh, S., 2003. Herd Behaviour and Cascading in Capital Markets: a Review and Synthesis. *European Financial Management* 9: 25-66.

Hirshleifer, D., Hong Teoh, S., 2010. *Psychological Influences on Financial Regulation and Policy*. In: Baker K. & Nofsinger J.R. (eds.). Behavioral Finance: Investors, Corporations, and Markets. New York: Wiley. 151-167.

Hirshleifer, D., Teoh, S.H., 2009. *Thought and behavior contagion in capital markets*. In: Hens T. & Schenk-Hoppé K.R. (eds.). Handbook of financial markets: Dynamics and evolution. Amsterdam: Elsevier. 2-46.

Hong, H., Kacperczyk, M., 2010. Competition and bias. *The Quarterly Journal of Economics* 125: 1683-1726.

Hong, H., Kubik, J.D., 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58: 313-351.

Hong, H., Kubik, J.D., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics* 31: 121-144.

Horowitz, N., 2011. *Art of the Deal: contemporary art in a global financial market*: Princeton University Press.

Howitt, P., 2002. Looking Inside the Labor Market: A Review Article. *Journal of Economic Literature* 40: 125-138.

Hribar, P., Jenkins, N.T., 2004. The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies* 9: 337-356.

Jackson, A.R., 2005. Trade Generation, Reputation, and Sell Side Analysts. *The Journal of Finance* 60: 673-717.

Janis, I.L., 1972. *Victims of Groupthink*. Boston: Houghton Mifflin.

Jegadeesh, N., Kim, W., 2010. Do analysts herd? An analysis of recommendations and market reactions. *Review of Financial Studies* 23: 901-937.

Kahneman, D., 1992. Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes* 51: 296-312.

Katz, M.L., Shapiro, C., 1985. Network Externalities, Competition, and Compatibility. *The American Economic Review* 75: 424-440.

Katz, R., Lazo, J., 2011. *The economic value of weather and climate forecasts*. In: Clements M.P. & Hendry D.F. (eds.). The Oxford Handbook of Economic Forecasting. Oxford: Oxford University Press.

Ke, B., Petroni, K., Safieddine, A., 1999. Ownership concentration and sensitivity of executive pay to accounting performance measures: Evidence from publicly and privately-held insurance companies. *Journal of Accounting and Economics* 28: 185-209.

Kelly, B., Ljungqvist, A., 2007. The value of research. *NYU Working Paper Series.*

Kemp, M.C., 1962. Economic forecasting when the subject of the forecast is influenced by the forecast. *The American Economic Review* 52: 492-496.

Keynes, J.M., 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.

Kim, A., Meschke, F., 2013. CEO Interviews on CNBC. *Working Paper (SSRN).*

Kindleberger, C., 1978. *Manias, Panics, and Crashes*. New York: Wiley.

King, M.F., Bruner, G.C., 2000. Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing* 17: 79-103.

Kirschenheiter, M., Melumad, N.D., 2002. Can Big Bath, and Earnings Smoothing Co-exist as Equilibrium Financial Reporting Strategies? *Journal of Accounting Research* 40: 761-796.

Kothari, S., 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105-231.

Krugman, P., 2007. Will there be a dollar crisis? *Economic Policy* 22: 435-467.

Kuran, T., Sunstein, C.R., 1999. Availability Cascades and Risk Regulation. *Stanford Law Review* 51: 683-684.

Kvale, S., 1996. *Interviews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks: SAGE.

Lacko, J.M., Pappalardo, J.K., 2010. The Failure and Promise of Mandated Consumer Mortgage Disclosures: Evidence from Qualitative Interviews and a Controlled Experiment with Mortgage Borrowers. *American Economic Review* 100: 516.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1992. The impact of institutional trading on stock prices. *Journal of Financial Economics* 32: 23-43.

Lazear, E.P., Shaw, K., Stanton, C., 2013. Making Do With Less: Why Productivity Rises During Recessions. *NBER Working paper.*

Le Bon, G., 1895. *The Crowd: A Study of the Popular Mind*. London: Ernst Benn.

Lehman Brothers, 2007. *Annual Report*.

Leibenstein, H., 1950. Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics* 64: 183-207.

Lim, T., 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56: 369-385.

Lin, H.-W., McNichols, M.F., 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101-127.

Littig, B., 2002. *Interviews mit Experten und Expertinnen*. In: Bogner A., Littig B. & Menz W. (eds.). Das Experteninterview. Berlin: Springer. 191-206.

Liu, X.G., Natarajan, R., 2012. The Effect of Financial Analysts' Strategic Behavior on Analysts' Forecast Dispersion. *The Accounting Review* 87: 2123-2149.

Ljungqvist, A., Marston, F., Wilhelm Jr, W.J., 2006. Competing for securities underwriting mandates: Banking relationships and analyst recommendations. *The Journal of Finance* 61: 301-340.

Longstaff, F.A., 2010. The subprime credit crisis and contagion in financial markets. *Journal of Financial Economics* 97: 436-450.

Lorenz, J., Rauhut, H., Schweitzer, F., Helbing, D., 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108: 9020-9025.

Lorge, I., Fox, D., Davitz, J., Brenner, M., 1958. A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. *Psychological Bulletin* 55: 337.

Lucas Jr, R.E., 1976. *Econometric policy evaluation: A critique*. Carnegie-Rochester conference series on public policy: 19-46. Manchester: North-Holland.

Machlup, F., 1955. The problem of verification in economics. *Southern Economic Journal*: 1-21.

Mackay, C., 1841. *Extraordinary Popular Delusions and the Madness of Crowds*. London: Office of the National Illustrated Library.

MacKenzie, D.A., 2006. *An engine, not a camera: how financial models shape markets*. Boston: MIT Press.

Malkiel, B.G., 1973. *A Random Walk Down Wall Street*. New York: Norton.

Malloy, C.J., 2005. The geography of equity analysis. *The Journal of Finance* 60: 719-755.

Malmendier, U., Shanthikumar, D.M., 2009. Do Security Analysts Speak In Two Tongues? *Working Paper University of Berkeley.*

Malmendier, U., Shanthikumar, D.M., 2007. Do security analysts speak in two tongues? *NBER Working paper.*

Marsden, P.V., Friedkin, N.E., 1993. Network studies of social influence. *Sociological Methods & Research* 22: 127-151.

Mathis, J., McAndrews, J., Rochet, J.-C., 2009. Rating the raters: are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics* 56: 657-674.

Matsumoto, D.A., 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483-514.

Mayring, P., 2000. Qualitative Content Analysis. *Forum Qualitative Social Research* 1: 1-11.

Mayring, P., 2003. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Beltz.

McAllister, I., Studlar, D.T., 1991. Bandwagon, underdog, or projection? Opinion polls and electoral choice in Britain, 1979-1987. *Journal of Politics* 53: 720-741.

McAndrew, C., Smith, J.L., Thompson, R., 2012. The impact of reserve prices on the perceived bias of expert appraisals of fine art. *Journal of Applied Econometrics* 27: 235-252.

McDonald, M.P., Thornburg, M.P., 2012. Interview Mode Effects The Case of Exit Polls and Early Voting. *Public Opinion Quarterly* 76: 326-349.

Mehran, H., Stulz, R.M., 2007. The economics of conflicts of interest in financial institutions. *Journal of Financial Economics* 85: 267-296.

Mei, J., Moses, M., 2005. Vested interest and biased price estimates: Evidence from an auction market. *The Journal of Finance* 60: 2409-2435.

Menkhoff, L., Taylor, M.P., 2007. The obstinate passion of foreign exchange professionals: technical analysis. *Journal of Economic Literature*: 936-972.

Merton, R.K., 1936. The unanticipated consequences of purposive social action. *American Sociological Review* 1: 894-904.

Merton, R.K., 1948. The self-fulfilling prophecy. *The Antioch Review* 8: 193-210.

Merton, R.K., Kendall, P.L., 1946. The Focused Interview. *The American Journal of Sociology* 51: 541-557.

Miceli, M.P., Near, J.P., Dworkin, T.M.U., 2008. *Whistle-blowing in organizations*: Routledge.

Michaely, R., Womack, K.L., 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12: 653-686.

Michaely, R., Womack, K.L., 2005. *Brokerage Recommendations: Stylized Characteristics, Market Responses, and Biases*. In: Thaler R.H. (ed.). Advances in Behavioral Finance. New York: Sage Foundation. 389-422.

Mikhail, M.B., Walther, B.R., Willis, R.H., 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74: 185-200.

Milgrom, P.R., Weber, R.J., 1982. A theory of auctions and competitive bidding. *Econometrica*: 1089-1122.

Milliken, F.J., Morrison, E.W., Hewlin, P.F., 2003. An Exploratory Study of Employee Silence: Issues that Employees Doní t Communicate Upward and Why. *Journal of Management Studies* 40: 1453-1476.

Morck, R., 2008. Behavioral finance in corporate governance: economics and ethics of the devil's advocate. *Journal of Management & Governance* 12: 179-200.

Morgenstern, O., 1928. *Wirtschaftsprognose*. Wien: Springer.

Morris, S., Shin, H.S., 2002. Social value of public information. *The American Economic Review* 92: 1521-1534.

Morrison, E.W., Milliken, F.J., 2000. Organizational silence: A barrier to change and development in a pluralistic world. *Academy of Management Review*: 706-725.

Morrison, E.W., Milliken, F.J., 2003. Speaking up, remaining silent: The dynamics of voice and silence in organizations. *Journal of Management Studies* 40: 1353-1358.

Morwitz, V.G., Pluzinski, C., 1996. Do polls reflect opinions or do opinions reflect polls? The impact of political polling on voters' expectations, preferences, and behavior. *Journal of Consumer Research*: 53-67.

Muth, J.F., 1961. Rational expectations and the theory of price movements. *Econometrica*: 315-335.

Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69: 1519-1554.

Noelle-Neumann, E., 1974. The Spiral of Silence: A theory of public opinion. *Journal of Communication* 24: 43-51.

Noelle-Neumann, E., 1984. *The Spiral of Silence: Public Opinion - Our Social Skin*. Chigago: Chigago University Press.

Nofsinger, J.R., Sias, R.W., 1999. Herding and Feedback Trading by Institutional and Individual Investors. *The Journal of Finance* 54: 2263-2295.

North, D., 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

O'Brien, P.C., 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research*: 286-304.

O'Brien, P.C., McNichols, M.F., Hsiou Wei, L., 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43: 623-650.

Oeller, L.-E., Barot, B., 2000. The accuracy of European growth and inflation forecasts. *International Journal of Forecasting* 16: 293-315.

Parker, L.E., 1993. When to fix it and when to leave: Relationships among perceived control, self-efficacy, dissent, and exit. *Journal of Applied Psychology* 78: 949.

Pfadenhauer, M., 2009. *At Eye Level: The Expert Interview - A Talk between Expert and Quasi-expert*. Houndmills New York: Palgrave/MacMillan.

Phillips, D.J., Zuckerman, E.W., 2001. Middle-Status Conformity: Theoretical Restatement and Empirical Demonstration in Two Markets. *American Journal of Sociology* 107: 379-429.

Pierdzioch, C., Ruelke, J.-C., Stadtmann, G., 2012. A Note on Forecasting Emerging Market Exchange Rates: Evidence of Anti-herding? *Review of International Economics* 20: 974-984.

Premeaux, S.F., Bedeian, A.G., 2003. Breaking the Silence: The Moderating Effects of Self Monitoring in Predicting Speaking Up in the Workplace. *Journal of Management Studies* 40: 1537-1562.

Prendergast, C., 1993. A Theory of" Yes Men". *The American Economic Review* 83: 757-770.

Prendergast, C., Stole, L., 1996. Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning. *Journal of Political Economy*: 1105-1134.

Rajan, R., Wulf, J., 2006. The flattening firm: Evidence on the changing nature of firm hierarchies from panel data. *Review of Economics and Statistics* 88: 759-773.

Rajan, R., Zingales, L., 2001. The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms. *Quarterly Journal of Economics* 116: 805-851.

Rajan, R.G., 1994. Why bank credit policies fluctuate: a theory and some evidence. *The Quarterly Journal of Economics* 109: 399-441.

Rajan, R.G., 2012. Presidential Address: The Corporation in Finance. *The Journal of Finance* 67: 1173-1217.

Raman, K., 1981. Municipal Financial Reporting: Managing the Numbers. *Public Budgeting & Finance* 1: 56-61.

Reinhart, C.M., Rogoff, K., 2009. *This time is different: Eight centuries of financial folly*: Princeton University Press.

Renneboog, L., Spaenjers, C., 2013. Buying beauty: On prices and returns in the art market. *Management Science* 59: 36-53.

Repenning, N.P., Henderson, R.M., 2010. Making the Numbers? Short Termism, and the Puzzle of Only Occasional Disaster. *NBER Working paper.*

Robert III, H.M., Robert, H.M., Robert, S.C., Evans, W.J., Honemann, D.H., Balch, T.J., 2011. *Robert's rules of order: Newly revised*: Da Capo Press.

Robinson, E., 2009. Research Renegades. *Bloomberg Markets Magazine*, October 9, 2009: 32-44.

Rost, K., Osterloh, M., 2009. Management fashion pay-for-performance for CEOs. *Schmalenbach Business Review* 61: 119-149.

Rothschild, K.W., 1964. Cobweb cycles and partially correct forecasting. *The Journal of Political Economy* 72: 300-305.

Roubini, N., Setser, B., 2004. The United States as a Debtor Nation: The Sustainability of the US External Imbalances. *Draft. New York University, New York.*

Ryan, K.D., Oestreich, D.K., 1991. *Driving fear out of the workplace: How to overcome the invisible barriers to quality, productivity, and innovation*. San Francisco: Jossey-Bass.

Salganik, M.J., Dodds, P.S., Watts, D.J., 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311: 854-856.

Samuelson, P.A., 1965. Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review* 6

Samuelson, P.A., 1974. Challenge to judgment. *The Journal of Portfolio Management* 1: 17-19.

Sarno, L., Taylor, M.P., 2001. Official intervention in the foreign exchange market: is it effective and, if so, how does it work? *Journal of Economic Literature* 39: 839-868.

Scharfstein, D.S., Stein, J.C., 1990. Herd Behavior and Investment. *The American Economic Review* 80: 465-479.

Schooler, J.W., Ariely, D., Loewenstein, G., 2003. The pursuit and assessment of happiness can be self-defeating. *The Psychology of Economic Decisions* 1: 41-70.

Schulz-Hardt, S., Brodbeck, F.C., Mojzisch, A., Kerschreiter, R., Frey, D., 2006. Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology* 91: 1080-1098.

Shiller, R.J., 1995. Conversation, information, and herd behavior. *The American Economic Review* 85: 181-185.

Shiller, R.J., 2005. *Irrational exuberance (2^{nd})*: Random House Digital, Inc.

Silver, N., 2012. *The Signal and the Noise: Why So Many Predictions Fail-but Some Don't*: Penguin Press.

Simon, H.A., 1954. Bandwagon and underdog effects and the possibility of election predictions. *Public Opinion Quarterly* 18: 245-253.

Simpson, A., forthcoming. Does Investor Sentiment Affect Earnings Management? *Journal of Business Finance & Accounting*

Sinha, P., Brown, L.D., Das, S., 1997. A Re-Examination of Financial Analysts' Differential Earnings Forecast Accuracy. *Contemporary Accounting Research* 14: 1-42.

Sockin, M., Xiong, W., 2013. Feedback effects of commodity futures prices. *NBER Working paper*.

Soros, G., 1987. *The alchemy of finance: Reading the mind of the market*. New York: Wiley.

Stanley, J.D., 1981. Dissent in Organizations. *Academy of Management Review*: 13-19.

Stein, J.C., 2002. Information production and capital allocation: Decentralized versus hierarchical firms. *The Journal of Finance* 57: 1891-1921.

Strasser, G., Titus, W., 1985. Pooling of unshared information in group decision making. *Journal of Personality and Social Psychology* 57

Sudman, S., 1986. Do exit polls influence voting behavior? *Public Opinion Quarterly* 50: 331-339.

Surowiecki, J., 2005. *The wisdom of crowds.* New York: Anchor.

Tajfel, H., Turner, J.C., 1986. *The social identity theory of inter-group behavior*. In: Austin W.G. & Worchel S. (eds.). Psychology of intergroup relations. Burnham: Nelson-Hall Publishers.

Tangirala, S., Ramanujam, R., 2008. Exploring nonlinearity in employee voice: The effects of personal control and organizational identification. *The Academy of Management Journal* 51: 1189-1203.

Temin, P., Voth, H.-J., 2004. Riding the south sea bubble. *American Economic Review* 94

The Economist, 2009. Please do feed the bears. *The Economist*, October 3, 2009: 16-16.

The Economist, 2013. Missing the mat. *The Economist*, July 20, 2013: 86-87.

Toonkel, J., 2012. *Morgan Stanley's Facebook analyst: sober man in world of hype*. Reuters, May 31, 2012.

Topchik, G.S., 2007. *The first-time manager's guide to team building*. Washington: AMACOM.

Trueman, B., 1994. Analyst forecasts and herding behavior. *Review of Financial Studies* 7: 97-124.

Vahey, S., Wakerly, L., 2013. Moving towards probability forecasting. *BIS Working Papers* 70.

Veblen, T., 1899. *The Theory of the Leisure Class: An Economic Study of Institutions*. New York: Dover.

Venn, J., 1866. *Logic of Chance*. Lond and Cambridge: MacMillan and Co.

Vogel, L., 2007. *How do the OECD Growth Projections for the G7 Economies Perform?: A Post-Mortem*. Paris: OECD Publishing.

Welch, I., 1992. Sequential sales, learning, and cascades. *The Journal of Finance* 47: 695-732.

Welch, I., 2000. Herding among security analysts. *Journal of Financial Economics* 58: 369-396.

Wenzelburger, J., 2006. *Learning in economic systems with expectations feedback.* Berlin: Springer.

Wermers, R., 1995. *Herding, Trade Reversals, and Cascading by Institutional Investors*. Working Paper, University of Colorado.

Westphal, J.D., Clement, M.B., 2008. Sociopolitical dynamics in relations between top managers and security analysts: favor rendering, reciprocity, and analyst stock recommendations. *The Academy of Management Journal* 51: 873-897.

Wilcox, C., 2010. *Groupthink: An Impediment to Success*: Xlibris Corporation.

Williamson, O., 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

Withey, M., Cooper, W., 1989. Predicting exit, voice, loyalty, and neglect. *Administrative Science Quarterly* 34: 1-539.

Womack, K.L., 1996. Do brokerage analysts' recommendations have investment value? *The Journal of Finance* 51: 137-167.

Zarnowitz, V., 1992. *Facts and Factors in the Modern Evolution of US Economic Fluctuations*. In: Zarnowitz V. (ed.). Business Cycles: Theory, History, Indicators, and Forecasting. Chigago: University of Chicago Press. 77-124.

Zwiebel, J., 1995. Corporate Conservatism and Relative Compensation. *The Journal of Political Economy* 103: 1-25.

Appendix

Appendix A to Chapter 3

Appendix A1: Interview Procedure and Data Analysis

To generate a systematic and replicable qualitative analysis, the in-depth interviews were conducted following the method of "focused interviews" (Merton and Kendall 1946, p. 541). Focused interviews are conducted as individual, face-to-face interviews in which the interviewer poses questions in accordance with a previously designed interview guide. In our case, this guide was based on the propositions which are derived from the theories on herding behavior and closely related areas reviewed in Chapter 2 and 3. The interview guide is provided in Appendix A.2. In the literature, the term focused interview is often used interchangeably with the term "semistructured interview" (see, e.g., Merton and Kendal 1946, p. 546; Kvale 1996, p. 5). Both terms imply that the interviewer uses an interview guide to pose the questions in a previously defined form while the question sequence can be varied according to the interview situation and the respondent can answer with free-form responses (Glaeser and Laudel 2006, p. 39).

There are four reasons why we decided to employ a qualitative study using focused interviews. Firstly, Merton and Kendall (1946, p. 541) suggest the use of focused interviews because "the array of reported responses to this situation enables the investigator to test the validity of hypotheses derived from […] social psychological theory", which is similar to our context here. Secondly, we refrained from sending questionnaires to organizations because it is unclear whether an addressee responds personally to the questions or one of his or her subordinates, in particular if the addressee is of high rank. Thirdly, as we aimed to interview a broad range of people in diverse positions and organizations, it was important to have the flexibility to react to the respondents' answers and adapt to the diverse situations they described. This is only possible in face-to-face interviews, where the interviewer can pose an additional question to clarify a narrated situation (see, e.g., Bewley 1998, p. 472). Fourthly, due to this flexibility and the meeting with the interviewee, the interviewer gains additional information about the environment the respondent is working in, which helps to correctly compare and link the statements provided across the various interviews (see, e.g., Helper 2000; Diekman 2001, p. 371-381).

Although research in economics is usually seen as a strictly quantitative matter, there are prominent examples in which economists have employed interview surveys to explore a phenomenon, test the external validity of an economic concept and further economic theory. Famous examples include Blinder and Choi's (1990) and later Bewley's (1999) interview surveys on sticky wages in labor economics, an area well known for its quantitative empirical rigor. More recent examples of interview surveys are Bloom and Van Reenen's (2007; 2010) analysis of management practices or Lacko and Pappalardo's (2010) work on the use of mandated consumer mortgage disclosures.

Sample selection

Bewley (1998) clearly stated that his sample is far from random, because many randomly selected, possible respondents refused to give an interview due to time constraints or other demurs. He explains that "the best information [to a find an interview partner] was word of mouth […], and a nearly sure way to obtain an interview was a personal contact who was trusted by the respondent […]. I therefore arranged as many interviews as possible through family, friends, and acquaintances, including people I had interviewed, always targeting the categories of firms I had in mind." (Bewley 1998, p. 473). We did exactly the same when arranging our interviews. We are left with a sample that is non-random but structured in the dimensions of the size of the financial services providers (number of employees) and the rank and position of the respondent. Thus, we are able to provide a comprehensive, balanced and insightful sample.

Interview procedure and data analysis

All but three of the focused interviews were conducted in the respondents' offices or meeting rooms of their employers, financial providers located in Zurich, Switzerland. One of the remaining interviews took place in the respondent's home and two of them in the interviewers' office at the University of Zurich. The interviews were conducted in German in accordance with the interview guide (see Appendix A2) They were digitally recorded and literally transcribed by the author, resulting in more than 300 pages of interview text. One of the respondents refused to be recorded. In this instance, the interviewer took detailed notes during the interview.

The data analysis was then coded following Mayring's (2003, p. 92-99) "scaled and structured content analysis". In order to present the analysis in a reproducible form, each step is shortly described. As Mayring's book (2003) is held in German, the terms are translated in accordance to Mayring's earlier summary in English (2000):

4. *Definition of the unit of analysis*: Interview transcripts of the twelve interviews.

5. *Definition of the dimensions of assessment:* The dimensions were defined as the degree of acceptance for a specific proposition. For example, the dimension of Proposition P2 – *The larger the organization is, the higher are the costs of non-herding* – is the degree to which the respondent accepts that the employees' costs of non-herding are increasing in the size of the organization (i.e. the number of employees).

6. *Definition of the scaled categories of dimensions:* For each of the dimensions (step 2.), a corresponding scale was defined. In general, the scale of the answers to the propositions was defined by five categories, namely *strong* a*ffirmation*, weak affirmation, *denial of influence* (respondent sees no influence of the proposed factor), *strong contrariwise effect*, or *weak contrariwise effect (respondent sees an influence in the opposite direction)*.

7. *Formulating the definitions and rules for coding and identification of typical examples from the interviews*: This step of the procedure shows how systematic and replicable the scaled and structured content analysis is. The coding rules enable a reproduction of the coding by any other researcher. The coding rules define the scales and categories by which the respondents answers are coded for each dimension. This procedure is also supported by "prototypical statements" (Mayring 2000, p. 3), which are extracted from the interviews and used as an appropriate example for a specific coding rule. An example of a coding rule is provided in Appendix A3.

8. *First filtering of the data and marking of important text passages*: The transcripts have been filtered and all test passages that were related to any of the dimensions (step 2.) have been marked.

9. *Second filtering of the data and coding of the important text passages:* All the marked text passages from step 5. are coded in accordance with the coding rules.

10. *Revision and adaptation of the coding rules:* After the coding of about fifty percent of the transcripts, the coding rules are revised and, if necessary, adapted to arrive at a stricter and more systematic and comprehensible set of coding rules. This incremental process increases the reproducibility of the coding results (see, Mayring 2003, p. 53).

11. *Final coding and quantitative analysis of the results:* On the basis of the revised coding rules, a final coding is made for all the marked text passages. All the coded passages of an interview that have been assigned to the same dimension are quantitatively assessed by identifying the most often answers to a certain proposition. This leads to the final classification if the respondent generally affirms or denies a proposition. It is important that this technique uses all text passages that relate to a particular dimension (i.e. proposition) and not only the passages that follow directly after the question (concerning the proposition) was posed. This helps to identify the reliability of the respondents' answers. Contradictory statements are classified as undecided.

Generalizability of results and possible biases

The number of interviews was defined on the basis of the different categories of financial service providers and of the functions and positions the respondents are working in. Considering also the enormous effort involved in finding a respondent, conducting the interview, transcribing the tape and coding the data, twelve interviews seemed to be a reasonable number. "In current [psychological] interview studies, the number of interviews tend to be around 15 plus-minus 10" (Kvale 1996, p. 102). As this interview study analyzes propositions derived from sociopsychological theories, we can certainly compare its scope with the scope of psychological studies. Kvale (1996, p. 102-103) argues further that small numbers of interviews can also provide valuable insights into tendencies and mechanisms of human behavior. In economics, the picture is similar. For example Blinder and Choi (1990) conducted 19 interviews in their study on economic theories of wage stickiness. Hence, based on this reasoning, we think that we are able to use the results of our interview survey to provide tendencies in difference and congruence between theory and practice on human herding behavior.

There are several possible biases when using interviews to test propositions. As an interview is always conducted in a certain social environment, it can be biased by many different factors, such as age, sex, status, and group membership among others (see, e.g., Littig 2002; Pfadenhauer 2009). The social environment can trigger a social desirability bias (King and Bruner 2000). In addition, Howitt (2002) states that economists are generally skeptical about results from interview surveys. He uses the example of Bewley's (1999) interview study on sticky wages and names three major drawbacks. First, Bewley's interviews are not structured. Second, Bewley's interview analysis is not structured either. Third, Bewley asked his respondents not only about what they did but also what they think they would do. Our interview study does not have such drawbacks. In contrast to Bewley, we employed focused, semistructured interviews, used a highly structured, reproducible method for the analysis and did not ask the respondents about what they think, but only about their experience. However, while we have to acknowledge that our results might also be biased by social desirability, we argue that the bias is low, due to three facts. Firstly, the questions did not address issues of high moral standards, nor were they embarrassing for the respondents. In addition, the questions are formulated in an objective and impartial form which does not suggest any desirability, i.e. that a particular kind of response is preferable (see Diekmann 2001, p. 391-399). Secondly, the interviews are conducted in a business context, in which the respondents are used to the question-and-answer structure of an interview. Thirdly, the respondents did not have to fear any consequences as we, the interviewers, did not know their supervisors, are not involved in their business network, and were able to assure full anonymity. In addition, two of them were already retired, implying even weaker incentives to whitewash the issues in question.

Appendix A2: Interview guide

Dissertationsprojekt: “The Costs of Non-Herding”
Von Reto Cueni
Lerhstuhl von Prof. Dr. Bruno S. Frey

Interviewleitfragen ***(Please see below for the English translation)***

→ *Dieses Interview findet im Rahmen meines Dissertationsprojektes statt. Sämtliche Aufzeichnungen werden von mir hoch vertraulich behandelt; die Interviews werden nur anonymisiert weiterverwendet.*
Einleitende Fragen: E1: Was würde Sie unter “Herdenverhalten auf den Finanzmärkten” verstehen? E2: Kommt Ihrer Meinung nach solches Verhalten tatsächlich vor und wenn ja, wie entsteht es?
→ Überleiten auf “Herdenverhalten in Unternehmungen”
F1: Sind Ihnen Situationen bekannt, in denen Mitarbeitende bei einer Entscheidungsfindung (in Meetings, Gremien oder ähnlich) eigene Bedenken und kritische Meinungen nicht eingebracht haben?
→*Wenn ja, weiter:* *(Wenn Nein →F7)*
F2: Haben diese Mitarbeitenden dann dadurch eine Entscheidung

	begünstigt, die diese selbst eigentlich als falsch eingeschätzt haben?
F3:	Gab es Situationen, die wegen der **nicht geäusserten Bedenken** von Mitarbeitenden und somit aufgrund dieser **fehlenden Informationen** zu **falschen Entscheidungen** geführt haben?
F4:	Wie würden Sie solche Situationen beschreiben? Gibt es spezifische Merkmale? *(→ spezifische Merkmale abfragen wenn unklare Antwort)*
F5:	Wie würden Sie ein solches **Verhalten von eigentlich kritischen**, aber schweigenden **Mitarbeitenden erklären**? *(Spezifische Gründe? → Wenn befraget Person ähnliche Begriffe wie Kosten nennt (Risiko, Angst ausgestossen zu werden, etc.) für Klärung nachfragen*
F6:	(Wenn nicht schon erwähnt in F4) Sind Ihnen Situationen bekannt, in denen solche fehlenden kritischen Stimmen oder falschen Entscheidungen es den kritischen Mitarbeitenden immer weiter erschwert haben ihre gegensätzlichen Meinungen zu äussern und sich dadurch die falschen Entscheidungen immer weiter verstärkt haben? *(Kettenreaktion, Verstärkung → Herdenverhalten)*
F7:	→*(Wenn nein bei F1)* Denken Sie es gibt Gründe dafür, dass das oben beschriebene Verhalten in Ihrem Umfeld (Organisation) nicht vorkommt? *(→Gibt es vielleicht „ähnliche" Situationen bei Ihnen → würden Sie diese aber eventuell anders beschreiben?)*
F8:	Welche Faktoren sind Ihrer Meinung nach entscheidend, dass die Mitarbeitenden in einer Unternehmung ihre persönlichen

	Standpunkte vertreten und kritische Meinungen äussern können? *(→ spezifische Faktoren!)*
F9:	Kennen Sie aufgrund Ihrer Erfahrung bestimmte Typen von Banken oder anderen Firmen in der Finanzwirtschaft in denen kritische Meinungen nicht oder nur sehr schwer geäussert werden können. → Was haben diese Banken für Merkmale ? *(→ spezifische Faktoren!)*
F10:	Kennen Sie aufgrund Ihrer Erfahrung bestimmte Typen von Mitarbeitenden, die besonders dazu neigen, ihre kritischen Meinungen nicht zu äussern in Diskussionen, Meetings, etc. (>in Situationen der Entscheidungsfindung) → Was haben diese Arbeitenden für Merkmale? *(→ spezifische Faktoren!)*
F11:	(Für die kommenden Fragen): Wie würden Sie unsere Determinanten (Gründe) beurteilen, dass in einer Bank kritische Stimmen eher geäussert werden? *(Bitte geben Sie an, ob die Richtung des Effektes für Sie stimmt und wie wichtig sie die Determinante einschätzen)*
F11.1: *Hierarchie*	
	Je weniger hierarchisch organisiert (weniger authoritär) eine Bank ist, desto eher können kritische Meinungn innerhalb ihrer Organisation geäussert werden.
F11.2: *Grösse der Bank*	
	Je kleiner die Bank, desto besser können kritische Stimmen geäussert werden. → gemessen an Anzahl Mitarbeiter
F11.3: *Zusammensetzung der Eigentümerschaft*	

Banken im privaten Besitz geben den Mitarbeitenden ein besseres Umfeld um sich kritisch zu äussern, als börsenkotierte Unternehmen
F11.4: *Personalwesen:*
F11.4.1: Lohnpolitik: Je weniger variable Lohnanteile (Boni) desto eher werden kritischen Meinungen geäussert.
F11.4.2: (Nachhaken) Kennen Sie Unternehmen, die eine Personal-Strategie verfolgen, indem sie gezielt Mitarbeitende auswählen, die eher dazu neigen sich kritisch zu äussern?
F11.5: *Risikomanagement:*
F11.5.1: Das Risikomanagement hat die institutionalisierte Aufgabe des "Advocatus Diaboli", darum interessiert mich dort spezifisch, welche Merkmale Sie in einer Bank als unabdingbar für das Risk Management einschätzen, um diese Rolle wahrzunehmen und auch kritische Meinungen zu äussern. Haben. Sie bestimmte Erfahrungen damit gemacht?
F11.5.2: Hat der Grad der Unabhängigkeit des Riskomanagements innerhalb der Organisation Einfluss auf den Schwierigkeitsgrad für Mitarbeitenden ihre kritischen Meinungen zu äussern? *(Das Riskomanagement könnte den Mitarbeitenden als Basis dienen, um ihre kritischen Meinungen zu äussern in einem Entscheidungsprozess.)*
F11.6: Haben Sie weitere Anmerkungen zu diesen Determinanten oder andere Ideen dafür?
F12 (Für die kommenden Fragen) Wie würden Sie unsere individuellen Determinanten beurteilen, dass die Mitarbeitenden einer Bank ihre kritische Stimmen eher äussern werden? Hier geht es mir um die

persönlichen Faktoren der Menschen: *(Bitte geben Sie an, ob die Richtung des Effektes für Sie stimmt und wie wichtig sie die Determinante einschätzen*
F12.1: Je besser ausgebildet desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.2: Je älter desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.3: Je erfahrener in der ausübenden Tätigkeit desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.4: Je höher in der Hierarchie desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.5: Je vermögender desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.6: Je länger in der Firma desto eher sagen die Mitarbeitenden ihre kritischen Meinungen.
F12.7: Haben Sie weitere Anmerkungen zu diesen Determinanten oder andere Ideen dafür?
F13: Haben Sie sonstige Anmerkungen? → zum Forschungsthema → zur Art meines Vorgehens → zu anderen, verwandten Themen → Was würde Sie besonders interssierten in diesem Themenfeld des Herdenverhaltens von Mitarbeitern und generell von Herdenverhalten auf Finanzmärkten?
Herzlichen Dank, dass Sie sich Zeit für meine Forschung nehmen!

(Please see below for the English translation)

Dissertation project: “The Costs of Non-Herding”
by Reto Cueni
Chair of Prof. Dr. Bruno S. Frey / University of Zurich

Interview guide   ***(Translated copy, see German version for original)***

→ *This interview is part of my dissertation project. All recordings are strictly confidential; the interviews will be used only in anonymous form.*
Introducing questions: E1: What do you think of when you hear about herding behavior on financial markets?   E2: From your point of view, does such behavior exist and, if yes, how does it originate?
→ skip to Herding behavior in organizations
F1: Are you aware of situations in which employees did not communicate their concerns or criticism during a decision-making process (in meetings, in boards, etc.)?
→*If YES, continue:* *(if NO →F7)*
F2: Did these employees support a decision which they privately perceived as a wrong one?
F3: Were there situations that led to wrong decisions because

	employees did not communicate their private information or their objections during the decision-making process?
F4:	How would you describe such situations? Are there specific characteristics or mechanisms? *(→ask for specifics if answers are unclear)*
F5:	How would you explain the employees' behavior of holding a critical opinion but not communicating their objections or critical private information)? *(Specific reasons? → If respondent uses terms similar to costs (risks, fear of being/becoming a loner etc.) ask for specification)*
F6:	(If not already mentioned in F4) Are you aware of situations in which the employee's silence at the beginning of the decision-making process started to increase the costs of non-herding, leading to a reinforcement of wrong opinions later in the process? *(Chain reaction, Reinforcement → Herding behavior)*
F7:*(→ If NO in F1):*	Do you know why the herding behavior described does not occur in your organization? *(→Maybe you are aware of similar situations, which you would describe differently)*
F8:	Which factors are important in creating an organization in which the employees can communicate their private information and objections? *(→ Ask for specific factors!)*
F9:	Are you aware of specific types of banks or other financial service providers whose employees cannot communicate their criticism or can only do so against high resistance ? *(→Ask for specific characteristics!)*

F10: Are you aware of specific types of employees who are particularly prone to remaining silent and not communicating criticism against a prevailing opinion during a decision-making process (in meetings, boards, etc.) *(Ask for specific characteristics!)*
F11: (In the following set of questions): How would you assess our institutional determinants of employees' difficulties when they want to communicate criticism in a bank or a financial service provider, in your experience? *(→ Please indicate if you agree on the direction of the effect and how strong you experienced the impact of the determinant)*
F11.1: *Hierarchy*
The flatter the hierarchy in an organization (more decentralized), the easier for the employees to communicate their criticism in a decision-making process
F11.2: *Bank size*
The smaller the organization, the easier for the employees to communicate their criticism in a decision-making process → Size measured as the number of employees
F11.3: *Ownership structure*
Privately held banks provide a better environment (make it easier) than publicly listed banks for the employees to communicate their criticism in a decision-making process.
F11.4: *Human Resource Management*
F11.4.1: Wage policy: The lower the part of the salary that is performance related, the easier for the employees to communicate their criticism in a decision-making process
F11.4.2: (Follow-up question) Are you aware that your company pursues a HR

strategy to select a special type of employees who are more prone to communicate critical private information and criticism in a decision-making process?
F11.5: *Risk Management:*
F11.5.1: The risk management plays the institutionalized role of a devils advocate in a bank or a financial service provider. Thus, I am interested in the characteristics that are important for the risk management to fulfill its role and to be able to communicate information against a prevailing opinion. What is your experience in this matter?
F11.5.2: Does the degree of independence of risk management in the organization also influence the degree of difficulties employees in communicating criticism in general? *(Because the risk management might provide employees with a basis for expressing a critical perspective during a decision-making process.)*
F11.6: Have you further ideas or suggestions regarding institutional determinants?
F12 (In the following set of questions): How would you assess our individual determinants of employees' difficulties if they want to communicate criticism in a bank or a financial service provider in accordance with your experience? *(→ Please indicate if you agree on the direction of the effect and how strongly you experienced the impact of the determinant)*
F12.1: The higher the employee's education, the easier for him or her to communicate criticism during a decision-making process.
F12.2: The older the employee, the easier for him or her to communicate criticism during a decision-making process.
F12.3: The more experienced the employee in a certain job, the easier for him or

her to communicate criticism during a decision-making process.
F12.4: The higher an employee's hierarchical position in the organization, the easier for him or her to communicate criticism during a decision-making process.
F12.5: The wealthier the employee, the easier for him or her to communicate criticism during a decision-making process.
F12.6: The higher the employee's tenure in an organization, the easier for him or her to communicate criticism during a decision-making process.
F12.7: Have you further ideas or suggestions regarding individual determinants?
F13: Have you other suggestions? → on the research topic… → on the interview procedure… → on related topics… → What would be your aim in an analysis of the topic of employees' herding behavior and herding behavior in financial markets in general?
Thank you very much for your time!

Appendix A3: Example of coding rules

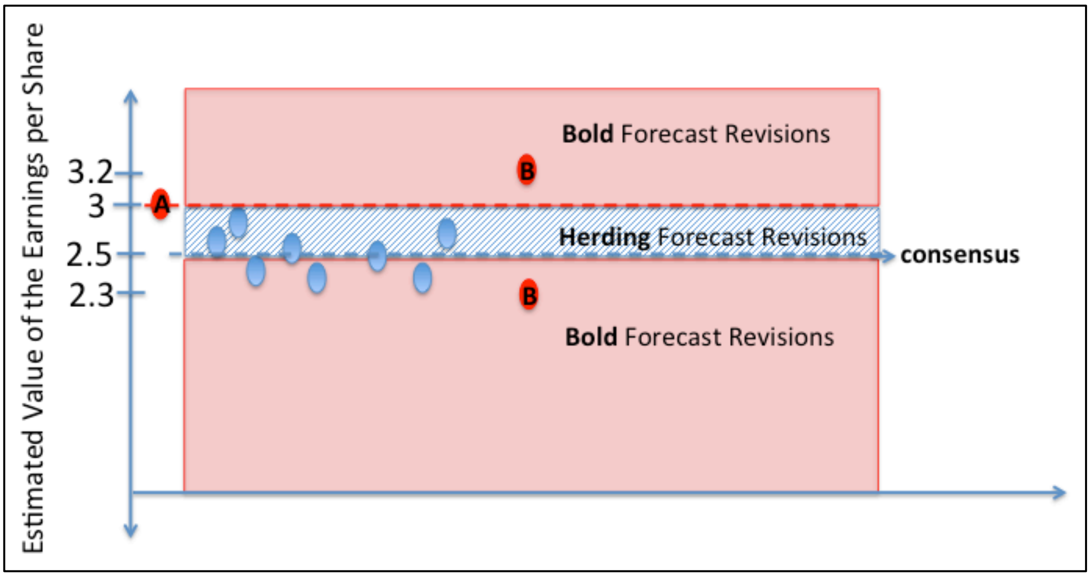
Kodierleitfaden - skalierend-strukturierende Inhaltsanalyse *(coding rules - scaled and structured content analyse)*					
Kodierung (nach betref. Hypothese) (*Coding in accordance with proposition*)	**Ausprägung/Kategorie *(Scale/category)***	**Definition *(definition)***	**Ankerbeispiel *(prototypical example)***	**Interviewnummer / Zeilennummer *(Number of the interview / of the line)***	**zusätzliche Kodierregel *(additional coding rules)***
Proposition P2	1 Starke Zustimmung *(strong affirmation)*	Grösse des Unternehmens hat starken Einfluss auf die CONH der MA *(Size of organization has strong influence on employees' CONH)*	Also sicher was das GEHÖR für kritische Meinungen angeht, da ist es sicher so, dass in kleineren Banken das gesagte eher Zuhörer findet. *("Surely, if it's about people listening more to critical voices in smaller banks, then for sure it is that way.")*	12/601	Keine *(None)*
	2 Schwache Zustimmung *(weak affirmation)*	Grösse des Unternehmens hat etwas Einfluss auf die CONH der MA *(Size of organization has some influence on employees' CONH)*	"Also ich denke das hat einerseits mit der grösse des Institutes zu tun, [...]". *("Well, I think that the size of the organization is somewhat linked [...].")*	10/433	Keine *(None)*
	3 Verneinung des Einflusses (Denial of influence)	Grösse des Unternehmens hat keinen Einfluss auf die CONH der MA *(Size of organization has no influence on employees' CONH)*	"Das ist keine Frage der Grösse, auch in grösseren Banken unterscheidet sich der tägliche Umgang nicht von kleineren." *("This is not a matter of size; normally even in big banks the daily work environment does not differ between a big bank and an institute of small or medium size.")*	11/505	Keine *(None)*
	4 Schwacher entgegengesetzer Effekt (weak contrariwise effect)	Grösse des Unternehmens hat schwachen entgegengesetzen Einfluss auf die CONH der MA *(Size of organization has strong contrariwise influence on employees' CONH)*	-kein Beispiel- *(-no example-)*	-	Keine *(None)*
	5 Starker entgegengesetzter Effekt (strong contrariwise effect)	Grösse des Unternehmens hat starken entgegengesetzer Einfluss auf die CONH der MA *(Size of organization has strong contrariwise influence on employees' CONH)*	-kein Beispiel- *(-no example-)*	-	Keine *(None)*

Appendix B to Chapter 4

Appendix B1: Additional Explanation for the Measure of Analysts' Herding Behavior

We follow Gleason and Lee (2003) and Clement and Tse (2005) in measuring the herding behavior of analysts' earnings estimates. The measure is illustrated in Figure B1. Imagine analyst *i* issues an EPS forecast on a stock *j*, pictured as estimate A (red dots) in Figure B1. Afterwards, many other analysts who cover the same stock *j* release their earnings estimates (blue dots). These estimates of the other analysts are used to generate the consensus. The revision of analyst *i* estimate (A) on stock *j* is depicted with two red dots (B). The binominal herding measure takes the value 1 if the forecast (B) of analyst *i* deviates both from his or her own most recent forecast (A) and from the consensus. All other estimates of analyst *i* that lie in the dashed area of the figure are defined as herding forecasts.

Figure B1: Additional Figure to Explain Herding Measure



Appendix B2:

Table B2: Additional Descriptive Statistics for Financial Coverage Subsample

Panel E: Financial Coverage (Privately held)
(N=5,417)

Variables	Mean	Min	25th	Median	75th	Max
Bold	0.78	0	1	1	1	1
BrokerSize	20.35	1	11	20	27	61
LagAccuracy	0.69	0.00	0.50	0.82	0.97	1.00
DaysElapsed	15.65	1	2	6	20	90
Horizon	96.45	31	61	71	117	346
Frequency	4.44	2	3	4	5	21
GenExperience	5.13	0	3	4	7	23
FirmExperience	3.10	0	1	2	4	21
Companies	18.43	1	13	18	22	51
Industries	2.80	1	2	2	3	22
Lead_Underwriter	0.61	0	0	1	1	1
Co_Underwriter	0.19	0	0	0	0	1
Syndicate_Member	0.10	0	0	0	0	1
No_Underwriting	0.10	0	0	0	0	1

Panel F: Financial Coverage (Publicly listed)
(N=16,406)

Variables	Mean	Min	25th	Median	75th	Max
Bold	0.77	0	1	1	1	1
BrokerSize	71.59	1	30	55	113	229
LagAccuracy	0.72	0.00	0.57	0.84	0.97	1.00
DaysElapsed	12.42	1	1	5	15	90
Horizon	94.18	31	60	69	98	348
Frequency	4.58	2	3	4	6	21
GenExperience	6.07	0	3	5	8	26
FirmExperience	3.81	0	1	3	5	24
Companies	19.08	1	13	18	23	87
Industries	2.84	1	2	3	3	22
Lead_Underwriter	0.80	0	1	1	1	1
Co_Underwriter	0.11	0	0	0	0	1
Syndicate_Member	0.02	0	0	0	0	1
No_Underwriting	0.08	0	0	0	0	1

Notes: The table reports summary statistics for the full sample (Panel A), the subsample of analysts covering the financial sector (Panel B), the subsamples of analysts from privately held (Panel C) and publicly listed brokers (Panel D), and the subsamples of analysts covering the financial sector from privately held (Panel E) and publicly listed brokers (Panel F). We retain only the last forecast an analyst *i* is issuing on a specific stock *j* in year *t*. The variable *Bold* equals 1 if the analyst's revised estimate deviates from both his last estimate and the prerevision consensus estimate, and 0 otherwise (see also Figure 1). *Private* equals 1 if the broker is privately held, and 0 if the broker is either a listed company or owned by a listed company. *BrokerSize* is the number of analysts employed at a specific broker. *LagAccuracy* is analyst *i*'s last year accuracy on stock *j*. *DaysElapsed* is the number of days that are elapsed between analyst *i*'s forecast an the most recent forecast of another analyst. *Horizon* is the number of days analyst *i*'s forecast is issued before the end of the financial period. *Frequency* is the number of forecasts analyst *i* makes during the financial year *t* for a stock *j*. *FirmExperience* is the number of consecutive years analyst *i* covers stock *j*. *GenExperience* is the number of consecutive years in which analyst *i* filed at least one forecast in the Institutional Brokerage Estimation System (IBES) since 1983. *Companies* and *Industries* are the numbers of stocks and sectors, respectively, which are covered by analyst *i*. *Lead_Underwriter*, *Co_Manager*, and *Syndicate_Member* equal 1 if the brokerage house acts in year *t* as either a lead underwriter, a co-manager, or a syndicate member in a U.S. IPO or SEO. *No_Underwriting* equals 1 if the brokerage house is not engaged in underwriting activities in year *t*.

Data sources: I/B/E/S, SDC Platinum Database, own compilation based on information from FINRA, FFIEC, SEC, Bloomberg Businessweek, Lexis Nexis, Wikipedia and company home pages.

Appendix C to Chapter 6

Appendix C1:

Table C1: Regression of Forecast Error on Time (Years) for Three Different Markets

Dependent variable: ***Yearly Median of AFEP***			
Variables	Weather	Art	Finance
Year	-0.000058*** (-3.04)	0.0061 (1.54)	0.0047* (1.92)
Constant	0.120645*** (23.82)	-12.038 (1.50)	-9.3873* (-1.90)
Observations	12	12	12
Adj. R2	0.43	0.11	0.19

Notes: The table reports OLS coefficient estimates, and t-values in parentheses. The models Weather, Art, Finance correspond to the respective data samples. The regression model is always the estimation of the *Yearly Median of AFEP* on *Year* and a *constant.* The dependent variable *Yearly Median of AFEP* is calculated as the median absolute forecast error in percentage of the actual outcome per calender year of the respective sample for weather forecasts (temperature prediction), art forecasts (presale price estimate) or financial forecast (earnings per share estimate). Year is the 4-digit number of the respective year between 1999 and 2010. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Data sources: MeteoSwiss, I/B/E/S, www.artron.net